DOCUMENT RESUME

ED 227 510 CS 207 414

AUTHOR White, Edward M.
TITLE Comparison and Contrast: The 1979 California State University and Colleges Freshman English Equivalency Examination.
INSTITUTION California State Univ. and Colleges, Long Beach.; English Council of the California State Univ. and Colleges.
PUB DATE 80
NOTE 81p.; For related documents, see ED 114 825-826 and CS 207 410-415.
PUB TYPE Reports - Descriptive (141) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS College Freshmen; Educational Testing; *Equivalency Tests; *Essay Tests; Higher Education; Objective Tests; Scoring; State Programs; Student Placement; Test Format; Testing Programs; *Test Interpretation; Test Norms; Test Reliability; *Writing Evaluation; *Writing Research
IDENTIFIERS California

ABSTRACT

The seventh administration of the California State University and Colleges Freshman English Equivalency Examination, a test consisting of an objective examination graded by the Educational Testing Service and an essay examination scored by California State University and College faculty, is analyzed in this report. After a preface and short index of previous volumes, the first section of the report contains a summary and fact sheet describing the test, its results, and studies being undertaken to examine the results. The next section recounts the actual administration of the test. The section on the examination's two essay questions consists of three subsections, two containing a report from the question leader, the question itself, the directions to graders, and six sample student papers for each of the questions, and the third listing the essay readers. Section four presents a statistical report on the examination scores. The report's three appendixes contain the examination's budget, a budget summary, and a summary report of descriptive statistics. (JL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED227510

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

COMPARISON AND CONTRAST

THE 1979 CALIFORNIA STATE UNIVERSITY AND COLLEGES FRESHMAN ENGLISH EQUIVALENCY EXAMINATION

EDWARD M. WHITE
DIRECTOR OF THE PROJECT
FOR
THE CALIFORNIA STATE UNIVERSITY AND COLLEGES

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Linda Bunnell Jones
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS207414

This monograph is one in a series describing the results of projects implemented under the California State University and Colleges Fund for Innovation and Improvement in the Instructional Process.

1980 Trustees of The California State University and Colleges

PREFACE

This report describes the seventh administration of The California State University and Colleges Freshman English Equivalency Examination. The development and initial implementation of the program are fully described in the first volume of Comparison and Contrast, which describes the 1973 program. The further development of the program is described in Comparison and Contrast 1974-78. Since this 1979 volume will not repeat material already published in earlier reports, anyone seeking a full account of the development of the English Equivalency Examination should acquire the earlier volumes from the Office of the Chancellor, The California State University and Colleges, 400 Golden Shore, Long Beach, California 90802.

Following the Table of Contents in this volume is a list of materials published in earlier reports that may be of particular interest to those receiving this report, which will go into detail only when new problems or new solutions occurred in 1979. This report does contain the essay questions that were used in 1979, the criteria for scoring, and six sample papers at the different score levels for each question. It also contains descriptive statistics for the 1979 test, a statistical report, and norm figures. This volume is slimmer than past ones, since it does not contain the usual in-progress report of the 1973 follow-up study. That study is nearing completion, and will be published as a separate volume shortly; it will detail the subsequent performance on a wide range of criteria of the total accessible 1973 test group.

The English Equivalency Examination program has been made possible through the continued financial support of the Fund for Innovation and Improvement in the Instructional Process of the California State University and Colleges. This year for the seventh consecutive year all administrative, publication, and essay examination expenses have been covered by this fund. The continued support of the program by the California State University English Council, composed of English Department Chairs and freshman English coordinators from all 19 campuses of the system, provides the program with a sound professional center.

Finally, it is worthy of note that, as the teaching and measurement of writing skills becomes more and more important at the college and university level, the English Equivalency Examination has served to involve a substantial number of university-level English faculty in the preparation and scoring of an essay examination according to the highest professional and psychometric standards. The continuing dialogue among the professionals involved in this program on all CSUC campuses has had a substantial effect on the teaching of writing and the teachers of writing throughout the state of California. It may turn out that this "in-service training" byproduct of English Equivalency will be the most important educational impact of the program. At the same time, we should not minimize the effect of the program upon the substantial numbers of students who have directly benefited from it.

TABLE OF CONTENTS

		Page
	Preface	i
	Short Index of Previous Volumes	iii
I.	Summary and Fact Sheet	1
II.	Administration of the Test Program	8
III.	The Essay Test	10
	A. Question 1	11
	a. Report by Question Leader: Dr. Eileen Lothamer	11
	b. Question 1	13
	c. Directions to Graders	14
	d. Six Sample Student Papers	15
	B. Question 2	21
	a. Report by Question Leader: Dr. Rex Burbank	21
	b. Question 2	24
	c. Directions to Graders	25
	d. Six Sample Student Papers	26
	C. Table Chart/Reader Roster, 1979	31
IV.	Statistical Report: Dr. Robert Bradley	32
	APPENDICES	
	A. Budget	46
	B. Budget Summary: A Report to the CSUC English English Council	47
	C. Descriptive Statistics	49

SHORT INDEX OF PREVIOUS VOLUMES

COMPARISON AND CONTRAST 1973, 1974, 1975, 1976, 1977, 1978

1. Summary and Fact Sheet

1973: 1 - 13
1974: 1 - 4
1975: 1 - 4
1976: 3 - 6
1977: 1 - 4
1978: 1 - 8

2. The Objective Test

1973: 28 - 30
1974: 13

3. The Essay Test: Reports, Questions, Sample Responses

1973: 30 - 61
1974: 15 - 39
1975: 19 - 36
1976: 11 - 26
1977: 7 - 18
1978: 11 - 32

4. Descriptive Statistics and Statistician's Reports

1973: 79 - 93, 132 - 167
1974: VII - XX, xxiii - xxxvi
1975: 51 - 68
1976: 31 - 47, 84 - 97
1977: 20 - 28
1978: 33 - 48, 55 - 74

5. Norm Samples

1973: 51 - 61 (7 CSUC campuses + UCLA and UCR)
1974: 41 - 49, VII - XI (San Jose State, Stanford, UCR, UCSB, UCLA)
1975: 54 - 62 (San Francisco State; CSU, Long Beach)
1976: 32 - 38 (CSU Northridge; CSU, Chico)
1977: 21 - 32 (CSU Fullerton; CSU, Long Beach)
1978: 34 - 44 (CSU Northridge; CSU, Fresno; San Jose State)

6. Follow-up Studies

1975: 71 - 91
1976: 48 - 63
1977: 33 - 48

7. Reports, Appendices, and Charts

Equivalency Testing in College Freshman English
1973: 94 - 125

Budgets
1973: 126 - 131
1974: XXI
1975: 93
1976: 64
1977: 48
1978: 49

Registration Form
1974: 7
1975: 7 - 10

English Council Committee Reports
1974: 13
1975: 12 - 16, 17 - 18
1978: 53 - 54

Organizational Charts
1975: 94 - 98
1976: 65
1977: 19
1978: 32

Media and Student Questionnaires
1975: 99 - 100
1978: 50

Data Processing Procedures
1975: 37 - 50

Four Issues for Faculty in Equivalency Testing
1975: 101 - 111

The Working of a Controlled Essay Reading
1976: 68 - 75

Son of EEE: The English Placement Test
1977: 57 - 58

The 1977 English Equivalency Reliability Study
1977: 51 - 54

SUMMARY AND FACT SHEET

The 1979 English Equivalency Examination proceeded smoothly, following, in general, the precedents established during the first six administrations. The upturn in enrollment that began in 1978 continued this year. It would be encouraging to think that renewed attention to writing in the schools, as shown, for example, by A.B. 65 and the California Writing Project, has increased the number of students who can profit by the EEE. Another possible reason for the increase may be the extra attention the associate director of the program gave to ensuring statewide distribution of the test brochure. The EEE seems to have become an established part of California education, increasingly well-known.

Certainly, one major indication of the continued success of the program is the number of times it is mentioned as a model for new testing programs, both statewide and on individual campuses. With large-scale testing programs in English now underway in the CSUC and in many other areas of education, the English Equivalency model and experience has become an extremely important base upon which to develop new programs.

The following chart compares the number of applicants, and their success rates, across the seven years of the English Equivalency Examination program. It shows the decline in number of enrollments until last year, partially compensated for by the increasing percentage of passing scores. The fairly sharp decline in the passing rate this year (a drop of 9 per cent from 1978) is discussed below in the statistical report. It is worth noting here, however, that a major component of that decline is the lowest mean score for any EEE test group on the multiple-choice test. Since that test examines the ability to read, particularly to read literature, we need to face the grim possibility that renewed efforts in the schools to improve student writing may be at the expense of instruction in reading.

STUDENTS APPLYING FOR ENGLISH EQUIVALENCY EXAMINATION CREDIT 1973-79

	Number	Number Passing	Percent Passing	Semester Units Earned
1973	4,072	1,363	33.5	8,178
1974	3,639	1,036	28.5	6,216
1975	3,303	981	29.7	5,886
1976	2,866	1,093	38.1	6,558
1977	2,216	883	39.8	5,298
1978	2,826	1,165	41.3	6,990
1979	3,346	1,081	32.3	6,486
Seven-Year Total	22,268	7,662	34.2	45,612

The multiple-choice test continued to be the CLEP Subject Examination, Analysis and Interpretation of Literaure, developed and graded by the Educational Testing Service. The essay test was produced by two CSUC Test Development Committees and scored by The California State University and Colleges faculty at the controlled essay reading at Kellogg West on the campus of California State Polytechnic University, Pomona, June 16-19, 1979.

In addition to the test papers, readers also scored norm samples written by 197 students from two CSUC campuses (Humboldt State University and California Polytechnic State University, San Luis Obispo).

Dr. Robert Bradley, test officer at California State Polytechnic University, Pomona, was once again in charge of the statistical operation. As has been the case in the past, the controlled essay reading achieved an extremely high rate of consistency among the readers. The statistical report estimates the reading reliability of the total essay score at .90 (see Section IV below). The essay scoring followed closely the pattern of previous years. The four independent readings on a 6-point scale were added to yield a single essay test score on a scale from 0-24. Essay total scores of 14 or above when combined with passing scores on the machine-scored test allowed a student to pass; eighteen students with a total score of 13, but with objective scores above 60 (that is, above the 85th percentile on national norms) were also allowed to pass.

For the third year, a second level "placement" pass was established. Students passing the test at this level were exempted from the required CSUC English Placement Test, even though their scores were not high enough to deserve credit. Any student with a total essay score of 15 or higher received a placement pass, whatever the multiple-choice score; an essay score of 11 or higher earned this pass when combined with a multiple-choice score of 35 or higher. When we add the 1726 students (51.6%) who received this placement, or exemption, pass to the 1081 (32.3%) who received credit for freshman English through the English Equivalency Examination, approximately 84% of the 1979 EEE tests received some form of pass--which means that approximately 84% of the test applicants received some substantial benefit from the test.

Meanwhile, the follow-up study of the 1973 test group (which has been delayed because of Proposition 13 budget cuts) nears completion. When printed in 1981, this study should give definite answers to the questions a large-scale equivalency testing program raises for a state university system--in particular to the question of how students gaining credit differ in their course plans and grades from those not gaining credit. The staffing, financial, and educational implications of this study are obviously important.

The 1980 English Equivalency Examination will be offered on May 3, 1980 and the essay will once again be scored by CSUC faculty in the middle of June. Among the issues to be addressed in the 1979-80 academic year by the program directors and English Council committees are 1) a new study of the validity of the EEE in the light of changing curricula and student populations, 2) a review of the multiple-choice portion of the test, now being revised by the Educational Testing Service, and 3) a review of funding, in particular ETS charges, in the light of budgetary cut-backs as a result of the passage of Proposition 13.

1979 ENGLISH EQUIVALENCY EXAMINATION FACT SHEET

Test Date:	April 28, 1979
Test Used:	CLEP Subject Examination, Analysis and Interpretation of Literature (90 minutes) in combination with Two 45-minute essay questions prepared by CSUC English faculty (90 minutes).
Number of Candidates:	3,346
Number Passed:	1,081
Percentage Passed:	32.3%
Number Exempted from Placement Testing	1,726
Percentage Receiving Exemption Pass	51.6%
Number of Semester Credit Units Earned	6,486

Scoring Data:

Minimum Passing Scores for credit:	
Objective Part	45 (CLEP scale) achieved by 52.3%
Essay Part:	14 (out of a possible 24 points; equivalent to 45 on the CLEP scale) achieved by 47.4% (Eighteen additional students with scores of 13 and objective scores above 60--above the 85th percentile on national norms--also passed).
Combined Score:	94 (49 objective 45 essay, each on the CLEP scale) 34.8% achieved 49 or better on the objective test, and 47.4% achieved 45 or better on the essay test.

Minimum Passing Scores
for EPT Exemption:

Objective Part:	35
Essay Part:	11 (out of a possible 24 points; equivalent to 36 on the CLEP scale) (Fifteen additional students with essay scores of 15 or above but objective scores below 35 also passed for EPT exemption).

Mean and Standard Deviation:

Objective Part Mean:	45.807
Standard Deviation	9.052
Essay Part Mean:	43.790
Standard Deviation	8.520
Combined Score Mean:	89.597
Standard Deviation	15.659

Statistical Data

Correlation between CSUC Essay and CLEP Objective Tests:	.59

Reliability:

CLEP Test:	.88
Essay Test	.745±.155
Composite Score:	.885±.045

Essay Reading

Number of Readers	36
Number of Colleges Represented	All 19 CSUC campuses

Grading Scale: 6 point

Number of Readings per Test:	4 (2 independent readings for each question; additional readings to reconcile discrepant grades)
Total Essay Readings:	Approximately 15,000
Weighting and Scaling:	The essay score was converted by the equipercentile method to the CLEP scale (20-80), and the scores on both parts of the examination were added. Each part received equal weight.

COMPOSITE
ENGLISH EQUIVALENCY EXAMINATION SCORING DATA
1973-1978

	1973	1974	1975	1976	1977	1978	1979
Minimum Passing Scores							
Objective Part	45	45	45	45	45	45	45
Essay Part: Raw score on 0-24 scale (Scaled score on 20-80 scale)	13(48)	13(53)	13(46)	13(48)	14(43)	14(48)	14(45)
Combined Score	100	105	99	101	92	97	94
Mean and Standard Deviation							
Objective Part							
Mean	47.448	49.444	47.180	48.995	47.050	47.831	45.807
Standard Deviation	8.854	9.001	9.069	9.425	9.119	9.246	0.052
Essay Part							
Mean	47.916	48.620	45.291	49.447	44.804	46.979	43.790
Standard Deviation	8.402	9.171	8.693	9.107	8.753	9.194	8.520
Combined Score							
Mean	95.363	98.064	92.474	98.442	91.854	94.810	89.597
Standard Deviation	14.869	15.855	15.402	16.456	15.616	16.424	15.659
Statistical Data							
Correlation Between CSUC Essay and CLEP Objective Tests	.4777	.5225	.5037	.5768	.53	.58	.59
Reliability							
CLEP Test	.88	.88	.88	.88	.88	.88	.88
Essay Test	.72±.18	.68±.15	.69±.14	.87±.04	.70±.11	.745±.125	.745±.155
Composite Score	.88±.05	.85±.05	.86±.04	.87±.04	.865±.035	.88±.04	.885±.045

ADMINISTRATION OF THE TEST PROGRAM

Once again, Dr. Edward M. White, California State College, San Bernardino, served as director of the test program. Associate Director was Dr. Rex Burbank, San Jose State University. Coordination of test administration took place through the test office of California State Polytechnic University, Pomona, under the direction of Dr. Robert Bradley. A new test manual was prepared by Betsy Barlow of the Educational Testing Service in Princeton, New Jersey, and the administration of the test went smoothly.

No changes were made in the existing statistical or computer operation, which allowed the pass and no pass letters to be ready for mailing to all students who took the examination within 24 hours of the end of the essay grading process. For the third year, a second computer operation took place under the supervision of Dr. Bradley, to allow for a second-level pass, for those students not receiving credit but performing well enough to merit exemption from the English Placement Test. (See Comparison and Contrast, 1977, Appendix C, for a discussion of the relationship between the EEE and the EPT.)

This year the same procedures that have been used in recent years to develop the essay test questions were again adopted. The advances in essay test development that were begun in 1975 were an important contribution to the field of essay testing, and we discovered once again that the time, effort, and money that were put into the creation of each essay question led to questions that were excellent for their purpose and that could be scored with accuracy. A committee based in northern California was under the leadership of Professor Rex Burbank, San Jose State University. The other in the southern part of the state was under the direction of Professor Eileen Lothamer, California State University, Long Beach. Working with each committee chair were three faculty from three different CSUC campuses. Thus, representatives of eight different campuses were involved in developing and pre-testing the essay questions. These committees developed essay questions during the course of the fall term and assigned them to freshmen composition sections on their own campuses for pretesting.

Following pretesting, grading sessions were held by the committees; careful assessment of scoring reliability and grade distributions for each proposed question then took place. Both questions were then revised and once again pretested. By the time the essay questions were ready for printing, the committees were confident that the questions would accurately reflect student writing skill and that these questions contained no ambiguities or other problems that would lead to

difficulty in grading. Results of the grading session in June demonstrated the value of the essay test development procedure. Once again, the questions worked very well, most particularly in allowing excellent students to show how well they could perform without distractions or ambiguities in the questions.

One important conclusion from this procedure is that any large-scale testing program must be prepared to spend both time and money on development of essay test questions. No matter how excellent a question may appear to be to a group of English professors, it may only be considered to be a good essay question after it has been pretested on real students and the resulting essays have been read carefully by trained readers. The readers can then judge whether a full and accurate range of scores is being elicited by that question and whether reliable scoring of the essays is possible. Certainly the essay test development and the essay scoring procedure for the English Equivalency Examination have developed into a highly reliable form of testing with important implications not only for testing but for the development of writing assignments in general. It has become clear to many of the teachers involved in the program that some essay assignments given to students in classroom situations lack the clarity, specificity, and ease of scoring that the English Equivalency Examination essay questions possess.

Finally, the California State University English Council continues to maintain close professional supervision over the entire test program. English Council committees evaluate the aims and procedures of the program and continued to support the test program by votes affirming the recommendations of these committees. (See Comparison and Contrast, 1975, p. 12 and 1978 Appendix C.) An additional sign of Council concern and sponorship occurred at the fall 1979 meeting, where, after much debate, a new committee was set up to review the EEE, with particular regard to the 1979 freshman English curricula at the various CSUC campuses. The Council has suggested that campus English Departments use the fact that they sponsor the test to "improve relationships and communication between our departments and the high schools." Only through continued participation by the English Council will the test continue to have the confidence of the English faculties throughout the system and continue to deserve that confidence.

III

THE ESSAY TEST

The essay portion of the English Equivalency Examination consists of two 45-minute writing tasks. Each of the assigned topics is designed to examine a different set of writing skills, through a different kind of topic. Students thus write on two separate questions in two different rhetorical modes. Since the total essay score combines the four scores given to the two essays, the test allows some compensation for students who may write very well in one mode but not as well in another. However, both modes of writing are frequent in college composition classes, and skill in both is normally expected.

The correlation between the two questions has, over the years, been about .4; this rather low correlation supports the assumption that the two questions are examining related but different skills.

The first question is in the expressive/analytic mode. It asks the student to describe something from his or her personal experience, and to speculate in some way upon the meaning of whatever is described. Each year, a committee develops a specific topic following this pattern. This question examines the student's ability to describe something in precise language, to organize that description in orderly sentences and paragraphs, and to use some abstract reasoning in relation to that description.

The second question is in the expository/analytic mode. It asks the student to read with care two passages on roughly the same topic, and to compare and contrast them. Each year, a committee develops a specific topic following this pattern. This question examines the student's ability to read with comprehension, to discern similarities and differences between complex quotations, and to develop an organized essay discussing coherently these similarities and differences.

The sections that follow describe the evolution of the questions that were used in 1979, give the questions, the criteria used to score them, and, finally, sample papers at each discrete score level.

A. Question 1

a. The following report was prepared by Dr. Eileen Lothamer, Question Leader for Question 1.

Summary Report on Question 1:

Question 1 for the 1979 English Equivalency Examination was designed and pretested by a committee of four state university English department faculty members: Dan McLeod, San Diego; Agnes Yamada, Dominguez Hills; John White, Fullerton; Eileen Lothamer, Long Beach, chairman. Veteran readers, these four faculty members met at California State University, Long Beach, October 20, 1978, instructed to bring three sets of experience-based questions, classroom tested and scored prior to that meeting. During the first meeting, the committee reviewed the questions and scoring guides for the previous English Equivalency Examinations, checked the responses to the classroom-tested topics and selected two trial topics to be pretested at each of the campuses and scored (with tentative six-point guides) before the next meeting. Many of the twelve topics designed by the committee were excellent for composition courses which could provide a context for motivating students or in which subsequent class discussion might lead to other kinds of process awareness. But as always, the committee had to consider questions which provided fair access to all students, no matter what their backgrounds; questions which elicited a full range of response, though, we hoped, not many score-one papers; questions which permitted holistic scoring without reader bias or idiosyncratic student responses.

At the second meeting, March 2, 1979, the group scored twelve sets of classroom papers, the pretesting of the two selected questions, then chose the one question that elicited the fullest range of responses and scores, developed a tentative scoring guide, and selected sample papers to be used and evaluated initially by all readers at Kellogg West. At both meetings, all papers were read and scored by the whole committee so that the final decision between the two questions was made by consensus, with optimism. The committee tightened the structural elements in the writing assignments and attempted to establish a clear three/four distinction on the scoring guide--a troublesome area quite often. Our optimism was tested at the reading in June at Kellogg-West.

Equipped with samples from the pretesting, the question leader read through ten batches of "live"--actual test papers, June 15, the evening before the statewide reading began. From these three hundred papers a rather discouraging discovery emerged. Despite the obvious insistence in the opening paragraph that the student select "a particular test" and despite the overt injunctions in the writing assignments to describe the "test," explain the "test," evaluate the

"test"--quite a few students succumbed to Saturday-morning testing sublimity and generalized their way from childhood to the grave, or wrote warmly about tests of all kinds being good for moral fiber. Optimism began to wane. Though few such papers had surfaced during the extensive pre-testing, the veteran faculty design committee had anticipated the temptation so that the scoring guide had as one of its key distinctions for a score of three, "those which are general and superficial."

At the table leaders' morning meeting, Saturday, June 16, six samples were first scored without the scoring guide so that the leaders might check the inductive process by which the Design Committee had arrived at the guide suggested for use. Others of the "live" samples demonstrating the range of three-level papers were then used to highlight a potential problem we resolved to confront head-on in the first readers' training session of the afternoon. After scoring some twelve samples, table leaders then read through test booklets (batches of thirty each) during that morning session to continue scoring-guide checks.

The afternoon session that day began with six samples first used with the table leaders. Readers were asked to examine the demands of the question closely. They then scored the first three papers without guide, full room discussion following. After examination of the scoring guide, another set of three papers was scored for room consensus. The first six papers had been selected to demonstrate the one-six range and the problems possible with different "kinds" of three-level papers. Another set of three papers with clear upper-half distinctions followed a brief break. Table leaders began to give specific instructions at 3:30 and scoring of live batches began.

The following morning and at each subsequent beginning of a morning or afternoon session, the question leader and table leaders (Lennis Dunlap, Chico; Dan Buerger, San Jose; Larry Kramer, San Bernardino) conferred on samples to reiterate the concern with holistic reading and use of the scoring guide. In addition, table leaders began to select "live" samples for occasional group scoring at a table. The search for strong score five and six papers was pursued by all readers as a reminder to avoid the temptation to droop to the 3/4 level. As with previous English Equivalency Examinations, responses to Question 1 were scored by two readers, and, where discrepancies of two or more points occurred, by a third reader. Throughout the reading, the table leaders' key responsibility was to provide continual spot checks on their readers' scoring reliability. Table leader check sheets were submitted to the question leader who also checked papers to monitor the tendencies to score less than holistically or without adequate reference to the scoring guide. Such problems were refreshingly

few after the first whole reading day, Sunday, June 17. The table leaders and I, in a post-reading meeting, agreed that readers did score with careful reference to the scoring guide and that the test booklets received serious and close readings. We all expressed some reservations about the number of generalized papers received, some joy over the decreasing number of score one or two papers, and some concern that the crop as a whole seemed less fresh than others in the past. Next year's Question Design Committee stands warned. However, the rewardingly low discrepancy rate pleased us all.

Although some of the readers may have dreamed later of the Kellogg-West reading as the one particular test that influenced their June spirits, with only three non-response papers out of the whole number, question 1 quite clearly permitted student response. Given the impromptu nature and the forty-five minute pressure, the students demonstrated, in the upper-half papers, a considerable amount of skill. The samples selected for this publication in no way represent the many and varied papers at the five-six level. We chose those which might not threaten a high school student contemplating the examination to show that style, wit, or profundity are not the deciding factors in holistic scoring.

Perhaps the tone of the question set the tone of the responses. Very few students ventured into lightness; most instructed us quite seriously about the tests and challenges human beings confront. We were indeed challenged by their papers and hope that the following samples reflect some of the tests we met vicariously with our students.

b. QUESTION 1. The following question was distributed to all students.

You have 45 minutes to write on the following topic:

"Life is one long test and we keep taking it." You may not share that comedian's summary of life, but certainly we all confront mental and physical challenges as we move through life--skiing, hang-gliding, jogging, proficiency tests, classroom tests, emotional tests with family, friends, loved ones. Reflect for a moment and select a particular test which in some way seems meaningful in your life.

WRITING ASSIGNMENT:

1. Describe the "test".
2. Explain your reaction and your handling of the challenges the test posed.
3. Evaluate the impact, influence, meaning of the "test" in your subsequent life.

c. The following directions for scoring were distributed to all readers engaged in the grading of Question 1.

Students should be rewarded for what they do well in response to the question. Here students are set a three-fold task: to describe a particular test, to explain their reaction to and handling of the challenge offered by the test, and to evaluate the impact, influence and meaning of the test in their lives.

Note that the question calls for description, analysis, and evaluation. In some papers the analysis may be implied through description and narration, but such implication should be appropriate to the approach used by the student.

An extremely well-written response may be scored a point higher than it would be scored on the basis of content alone. A poorly written response may be scored a point lower.

POSSIBLE SCORES:

6 A superior response will address itself to all aspects of the question. Though it may have occasional faults, it will be well-organized detailed, and generally well-written.

5-4 These scores will be useful for a well-handled paper which is weak in some aspects of the superior response; e.g., it may slight one of the parts of the question; it may not be as clearly organized as the superior paper; it may have some minor grammatical inconsistencies. Otherwise, the paper should be competently written.

3 This score will be useful for the following kinds of papers:

-- those which are only descriptive/narrative;
-- those in which the language is overly clichéd;
-- those which are overly repetitious;
-- those which are general and superficial.

2 This score is to be used for papers which exhibit serious weaknesses in structure, syntax, diction, and/or development.

1 This score is to be used for papers which show very little understanding of the question or suggest incompetence in structure, syntax, and diction.

* Non-response papers or those which argue with or avoid the question should be given to the table leader.

d. The following student responses to Question 1 were sample papers used during the reading to illustrate the grades on the 6-point scale:

SCORE OF ONE

Responsibility

Responsibility is the premier test of Life. Those individuals that handle, control, and manage responsibility are those that pass the test. The unfortunate creatures that cannot cope with responsibility, the ones that are dominated by a simple task, are the ones who shall fail.

The management of responsibilities is one of the most forbiding of hurdles blocking the pursuit of happiness, if this hurdle is overtaken fulfillment of the individual is within reach, it is an attainable goal.

Those who falter at this important step can become engulfed in worries and anxieties to an unbearable degree, our society has as its basis winning, these is no mercy for those who fail, it is either sink or swim, no in between can be reached.

Responsibility must be faced and overcome in order to function successfully in this life. Those who manage pass the test.

SCORE OF TWO

The test is similiar to the tests that I have been taking all my life, except for the fact that they have gotten more complex through the years.

The test is based on reading a short story, and answering questions to see how well you understand what you have read, and how well your vocabulary is. It is a short test, or long depending on how one looks at it.

I came prepared for this type of test. Not saying that I knew all the answers, but that I came prepared for this type of test. When taking the test I was very relaxed in handling the material, but I must say it was a rather complex test for me. But I went on to try and do the best I could from the knowledge I already have.

I feel that I handled the test appropriately. Many of the questions I did not know the meaning of of the words but I did the best I could.

I felt that the test was very important to me, because when I get my score back, I can register for a English class that will help me if I score low, or I will know the next step to take if I score average.

So, therefore I know that in reality the test is helping me more than it is the person who is giving it to me.

SCORE OF THREE

Classroom Test

There are three main reasons why classroom tests are given. They let the student and teacher know how well the student is learning the material, if he is progressing in the subject, and if he is passing or failing the course.

The most important thing that classroom tests reveall is whether or not the student is learning the material. The test will indicate what the student has learned and what the student has not learned. In this respoect tests have been very helpful to me as a student. I study for classroom tests until I believe that I thoroughly know the material. But on many occassions classroom tests have pointed out to me those areas of the material that I thought I knew when in reality I only had a vague idea of it. After the test has pointed out areas that I am having trouble in, I can seek help. If it had not been for the test, I probably would not have realized that I was not fully understanding the material. So, in this respect tests have been very helpful to me as a student.

Another thing that classroom tests point out is the student's progress in the course. From the test the student is made aware of whether or not he is tying all the materials he has learned together. From this he can tell if he is progressing in the class. By taking a test, I have sometimes been made aware of the fact that I am not tying what I learned previously in the class to what I have just recently learned. I may have understood this past section very well but if I cannot tie it in with previous sections then I am not really progressing in the course. The test points out to me as a student which areas are hindering my progress and what I need to do about the situation.

The third thing that tests do is point out to the student his course standing By looking at his grade on the classroom test the student can tell if he is doing good, fair or

poor in the class. This has helped me as a student because if I did good on the classroom test I know that I studying long enough and hard enough. On the other hand, if I did poorly on the classroom test I am made aware of the fact that I am not doing very well and I need to do something to change that.

In conclusion, classroom tests are very helpful to me. They reveal to me how well I am learning the material, how good my progress is and the degree of my class standing which is shown by the grade that I received.

SCORE OF FOUR

A test which has proven to be very meaningful in my life is my recent separation from my parents. My father found employment in Saudi Arabia, so my parents are living there while my sister and I live here in San Diego with relatives. My first reaction to hearing of my father's job and plans was complete disbelief. I could not believe he really would want to go to a seemingly backward country like Saudi Arabia. I felt that he was deserting me. After my parents left, I felt lonely and very afraid of the responsibilities which had previously been on their shoulders and would now be on mine. I had to do many things for myself which I would have been afraid to do before. The responsibility of paying bills, making doctor appointments, and solving many problems I hadn't ever faced seemed impossible at first. It is still difficult, but I am able to handle things much better now than I was in the beginning.

This experience has been and, I am sure, will continue to be very important in my life. I have changed in many ways, learned new things, and seeing things differently than I used to. I have become more independent and responsible and less shy and afraid to look out for myself. I am very happy for this, because I am sure that when I actually am living on my own, I will be ready and able to handle most of problems I will face. This separation has also helped me to appreciate my parents more. I can see them better as people now, not just as mom and dad who make the rules and buy the food. I'm certain that this will improve our relationship, especially later in life. One other meaningful change that was caused by my parents' move is the subsequent move from New York to San Diego for my sister and me This has also changed me as a person, because I have had to start a new school, make new friends, adjust to a different climate and lifestyle, and generally just start all over again. Altogether, I am very greatful for the experience I have had because of this move, and I know that in the long run, I will be a better and happier person for having lived through them.

SCORE OF FIVE

One may confront many a test in life, but a test which involved facing a reality about one's self would seem to be more difficult than most. The test involved in facing a reality is basically either accepting this reality or trying to remedy the reality and change it so as one is more pleased with the result than with the beginning substance. The most significant and meaningful test that I have had to confront in my life was facing the fact that I had (and still partially have) a birth defect.

My birth defect never incapacitated me, only acted as a bump in the road. It never caused me to cry, but caused me unhappiness at times. I never had reason to cry out in pain, but I moaned and uttered slightly profane phrases when the pain called on me. My otherwise "perfect" body has a curved spine. This defect is called scoliosis. Two years ago I began to experience a sharp pain between my shoulder blades. Also I noticed while looking in the mirror that my shoulders were not on a parallel line--one was lower than the other.

This test proposed two alternative courses to the solving of its problem. A bone specialist informed me that corrective surgery (resulting in a straight back and a body cast) accompanined by bed rest for nine moths--an eternity to a socially active teen like myself, would solve the problem once and for all. The other alternative being a succession of seven repititious isometric excercises performed five days a week, twice daily would also eventually (no specifics were given) solve the problem. I was also instructed that if I chose the exercise route, I was limited in rigourous physical exercise and in sleeping positions (visions of bedroom gymnastics ran through my head!) I chose the physical route. And so my test began.

The test had a great impact on my life in that I had never been limited to any extent in my active life. But now, here I was setting aside time for slightly painful sessions of exercise and turning down offers to participate in tennis and snow skiing (two of my favorite passtimes.) The test gave meaning to the fact that some people (who before my discovery had sparked just a hint of sympathy and understanding) were rendered almost totally incapacitated. I now felt empathetic with handicapped people, not just sympathetic. The test also gave me a reason to find out if I really cared for my own well-being. Would I follow the doctor's instruction? My test has proved positive. My defect has proved itself as an ironically welcomed test in my life.

Today, I've almost corrected what Mother Nature once defected. I'm happy with the way I handled my test and with

the ways my life has changed since the test began. One thing that the entire experience has proved to me is that no one is "defective," only rendered with a defect which they must either mentally or physically overcome.

SCORE OF SIX

Throughout my life, I have encountered many obstacles that have tested my mental capacities and physical abilities. Each of these challenges has contributed to my personal sense of identity and taught me much about my own limitations. One incident in particular comes to mind that tested me both mentally and physically and has often come to serve as a basis for reflection when I have doubts as to my abilities to handle subsequent challenges.

The incident itself was a day I spent flyfishing alone, high in the Trinity Alps in Northern California. Early one morning my father dropped me off several miles upstream with the idea that I would spend the day fishing my way back down stream and meet him at the campsites. This was one of the first times I had been left truly to my own devices in such rugged and isolated country and I looked forward to the event with, I must admit, a touch of trepidation. I was, however, determined to return to camp with my kreel bulging with native rainbow trout so I set off down the mountain stream in search of a promising-looking pool.

With each bend in the river I surveyed the banks and vegetation to determine the correct fly to use as bait. I had been used to relying on my father for advice on these matters, so this was a relatively new experience and I enjoyed the feeling of independence. Without my father there I had to recall everything I knew about flyfishing in order to know just how and where to present each fly and how to cast so that it would fall precisely where I wanted it to.

Perhaps more important than these mental tests were the extremely demanding physical challenges the stream itself presented. In several places, I was forced to cross from side to side of the fast moving stream in order to gain access to a particular pool. I had to rely on balance to cross as the moss-covered stones on the bottom were slippery and the water often came up to my thighs. I knew one wrong step could send me bobbing down the rapids on my backside. The ruggedness of the terrain made it necessary for me to climb over huge boulders in order to make my way down stream. Many times I had to inch my way along with the fingertips, clinging to a rocky wall with the raging water several feet below me.

I met each of these mental and physical challenges with determination and five miles and a few blisters later, realized that I had made it. I had ten beautiful trout to show off as a result of my fishing skill and had passed every physical test the river had to

offer. The feeling of self-satisfaction and pride in my own accomplishment was much more rewarding than the trout dinner I had that night. Today, when I am faced with a new experience and I have doubts that I can handle it, I think back to that day when I was left to my own resources in a very demanding situation, and succeeded, and I am given the confidence I need to continue.

B. Question 2

a. The following report was prepared by Dr. Rex Burbank, Question Leader for Question 2.

Summary Report on Question 2:

Essay Question 2 for the 1979 English Equivalency Examination was developed by a faculty committee consisting of Professors Jack Conner, CSU Hayward; William Robinson, San Francisco State; Hortense Thronton, Sacramento State; and me as chairman. All of us are members of the English Departments of our respective universities.

The committee met at San Jose State University three times during the 1978-79 school year. Our procedure followed much the same pattern as has been followed since 1975, that is, each member developed four essay topics. Of the total of sixteen topics, we selected four at our first meeting for pretesting. Each of us arranged for one topic to be pretested at his/her own campus, and at the second meeting the committee read the four sets of papers, which were written by students who had already completed or were then near completion of a second term of freshman English. From the four batches, we selected the one we felt was the most successful for our purposes. In making our selection, we evaluated the pretested topics in accordance with the following criteria:

1. Their capability of being scored holistically.
2. Their capability of being scored on a six-point scale, with a six as the highest and a one as the lowest scores.
3. Their capability of prompting responses from all the student examinees--could all high school students taking the test be reasonably and fairly expected to write on the topic?
4. Their capability of producing responses that had a structural arrangement (specifically, comparison and contrast) that could provide a foundation for group scoring.

Two of the four topics seemed very good, but as in former years, we felt a need to make some small but important revisions in both the directions and the passages themselves in order to clarify the assignment and focus attention upon the aspects of the passages that would enable the students to make the best use of their talents and training. At the third meeting we found that one topic gave us just what we wanted, while the other still needed some further small revisions, so our choice was easy. We drafted a scoring guide and selected a dozen samples, which we reproduced immediately and which I took with me to the holistic reading at Kellogg West in June.

As in previous years, Question 2 this year was designed to suggest an organizing principle upon which the candidate could write a 45-minute essay. It included two quotations, one a prose passage by Loren Eiseley and the other a stanza from a poem by Hillare Belloc. The assignment was to write an essay explaining each writer's attitude towards the relationship between science and human happiness, indicating the degree to which characters in each passage agree or disagree and to which the authors themselves agree or disagree with the unspecified "some" in the Belloc stanza. The structure therefore called for analysis, comparison, and contrast. With such an organizational principle set forth by the assignment, the student was afforded a ready-made framework within which to develop his/her thoughts about the two passages. The scoring guide was constructed to enable readers to score all essays by the same qualitative standards and was designed as a brief, general description of the papers--from strongest to weakest--readers could expect; the samples illustrated the qualities specified for each of the six points of the scale in the scoring guide.

At nine a.m., Saturday, June 16, I met with the three table leaders assigned to Question 2. We went through the twelve samples and found that our scorings were substantially the same as the first ones assigned by the Test Development Committee. We were thus ready for the afternoon training session with the readers, who gathered at 1:00 p.m.

In the afternoon, the readers met with the table leaders and me and discussed the question, how it was developed, and the basis upon which it was to be scored. They read and scored samples of 1, 2, 4, 4, 5 and 6 papers, and I explained the rationale of the scoring and referred them to the scoring guide. They read and scored more samples and discussed them among themselves at the tables. At 3:00 p.m. I asked the table leaders if they felt their readers were ready to begin "live" readings; all three said yes, and so the reading began. At 4:00 p.m. three more samples were distributed and read, and the reading day ended at five o'clock with a review of the scoring standards.

All together 28 samples were used during the reading. In addition to the twelve we started with, we located sixteen more from among the papers being checked by the table leaders. These we had xeroxed by the aides. We used fewer samples the second than the first day, and fewer the third than the second. Table leaders spent all their time checking the scores of readers at their table. Readers were asked to adjust their scoring in accordance with the scores given the samples by the table leaders. In addition to sampling, table leaders read and scored the papers themselves, and recorded both their own and the readers' scores on a "check sheet." The table leaders, who read and scored the papers without looking at the scores assigned by the readers, brought the checklist with the test booklets to me, and I then

read and scored them myself without looking at the scores given by the readers and table leaders. I was thus able to check continually the scoring of the readers and table leaders against my own. Where there were discrepancies, I could locate who was grading inaccurately and see that a correction was made; and where trends developed--such as a tendency to grade too high or too low or settle in the 3-4 range, the problem could be attacked immediately by further sampling and discussion of scoring standards. Using these methods, we achieved a high degree of agreement between first and second readers, and in about 95% of the cases the two readings produced scores that were within one point of each other. When a spread of two or more points occurred, the paper was read a third time, by one of the table leaders or by me. The number of papers with a spread of three points was very small, and those were radically uneven in quality, so that the discrepancies were understandable. I am satisfied that overall the papers were graded with a very high degree of uniformity, reliability and validity.

In scoring the essays we did not expect perfection from even the best student responses, but for a score of six we required that the paper show a high degree of skill in achieving what was asked in the assignment. The best demonstrated the student's ability to build on the implied organization with meaningful details, to develop the general points, and display a high degree of understanding of the quotations. Readers understood and took into account the fact that students were required to read two rather sophisticated passages and then to write meaningful commentaries on them in a matter of only 45 minutes--a demanding task; nevertheless, in order to be awarded the highest score, examinees were expected to demonstrate evidence of ability to perceive the complexities implicit in the quotations and to relate them in clear, standard written English.

Rex Burbank
San Jose State University
August 1979

b. Question 2. The following question was distributed to all students

Essay topic: read the following passages and write an essay according to the directions below.

This is a story about one of our great atomic physicists. This man, one of the chief architects of the atomic bomb, so the story runs, was out wandering in the woods one day with a friend when he came upon a small tortoise. Overcome with pleasurable excitement, he took up the tortoise and started home, thinking to surprise his children with it. After a few steps he paused and surveyed the tortoise doubtfully.

"What's the matter?" asked his friend.

Without responding, the great scientist slowly retraced his steps as precisely as possible, and gently set the turtle down upon the exact spot from which he had taken him up.

Then he turned solemnly to his friend. "It just struck me," he said, "that perhaps, for one man, I have tampered enough with the universe." He turned, and left the turtle to wander on its way.

--Loren Eiseley

"The path of life," some say, "is hard and rough
Only because we do not know enough.
When Science has discovered something more,
We shall be happier than we were before."

--Hilaire Belloc

Directions:

Explain each writer's attitude toward the relationship between science and human happiness.

To what extent do the scientist in Eiseley's passage and the "some" in Belloc's poem agree or disagree?

Do the two authors seem to agree more with the scientist or with the "some"?

c. The following directions for scoring were distributed to all readers engaged in the grading of Question 2.

The student is asked to write an essay that explains (1) each writer's attitude toward the relationship between science and human happiness; (2) the extent to which the scientist in the Eiseley passage and the "some" in the poem agree or disagree; and whether the authors seem to agree more with the scientist or with the "some". Papers should be scored for their overall quality, and the student should be rewarded for what she/he does well, but an essay that does not fulfill the three parts of the assignment should not ordinarily be given more than a 3.

An especially well-writen essay may be scored a point higher than it would be scored on the basis of content alone. A poorly written paper may be scored a point lower.

SCORES

6 A superior response will be a well-organized essay that addresses the three parts of the assignment. It will accurately explain each author's attitude--perceiving directly or implicitly the irony in the poem; show where the two selections are similar and how they differ; and indicate how the authors seem to agree more with the scientist than with the "some." An essay receiving a score of six will display a high degree of competence generally but may have slight flaws in writing. A 6 essay will support generalizations with appropriate details.

5 Essays in this category will address all parts of the assignmentt, display good overall interpretations of both passages, and clearly demonstrate competence, but they may be less fluent or thoroughly developed than the 6 papers, contain some inaccuracy in interpretation, or reveal minor grammatical errors or occasional awkwardness of expression.

4 Essays in this group will be generally well written but may lack the development of 6 and 5 papers, perceive similarities and differences somewhat less accurately, or have some errors in sentence construction or diction. They may deal with one aspect of the assignment by implication rather than by direct statement.

3 Essays in this category may show signs of clear writing but respond to only one or two parts of the assignment; contain misinterpretations of both passages; radically misinterpret one passage; show evidence of serious deficiencies in writing; or lack sufficient details to support their generalizations.

2 Essays assigned a 2 may radically misinterpret both passages; reveal serious faults in writing; display considerable irrelevance; be simplistic.

1 Essays given a 1 show almost no understanding of the question or passages, and the writing shows incompetence in sentence construction, usage and idiom.

Non-response papers and those completely off-topic should be given to the table leader.

d. The following student responses to Question 2 were sample papers used during the reading to illustrate the grades on the 6-point scale:

SCORE OF ONE

When the great atomic physicists saw the tortoise he was full of excitement. He had thought he could have had fun with it at home with his family. As soon as the tortoise stuck him he realized that it would be happier if it was left to wander its own way. Belloc says "We say the path to life is hard and rough only because we don't know." Personally I think Eiseley and Belloc agree to the fact that there is a lot of things we don't know, thus we say its rough or hard simply because we don't know. Sooner or later when things get discovered and we start knowing things life will be easier and we will be happier than before.

Both of the authors seem to agree on science. Eiseley feels that he has already tampered enough with the universe and doesn't feel like dealing with the tortoise. He'd rather leave it alone and let it wander whereas Belloc feels that it's only because enough things haven't been discovered so as to let us know what will make us happy or what is good from what is bad. In a way I agree with what Bellac says because the more things are discovered through science or tecnology the happier humans are because life gets to be easier and we wouldn't have to be wandering all the time. A solution or conclusion has been reached. The passage and the poem disagree at the point where the physicists chooses to let the tortoise wander the way it wishes to and Belloc says it's only because it doesn't know what is good or bad for it. The tortoise could be happy under the physicists care than it would be just out in the woods. So more discoveris have to be done to make even the animals as happy as humans. It seems like both authors look at human right scientifically and logically according to the way they feel or they define human rights in relation with science.

SCORE OF TWO

Attitudes of Loren Eiseley and Hilaire Belloc toward the relationship between science and human happiness are different in the following way.

Loren Eiseley, in his passage about one of our great atomic physisists, shows that this man who challenged nature in the most vialent way by taking the atomic power from it still thinks that there are some things that a man should not touch in order to be happy. This is described in the scientist's words when he was walking in the woods with his friend and picked up a tortoise but put it back after giving it some thought:"...perhaps, for one man, I have tempered enough with the universe."

Hilaire Belloc, however, describes in his passage about what "some" say about science a significantly different attitude. "Some" say that "...life is hard and rough only because we don't know enough" and that scientists should temper with nature as much as possible and that "when science has discovered something more, we shall be happier than we were before."

SCORE OF THREE

Loren Eiseley and Hilaire Belloc express different views toward the relationship between science and human happiness. Their attitudes voice the contravercy between the need for the knowledge which science offers and it's conscious limitation to avoid the possible destructive ends it sometimes promotes, such as the atomic bomb mentioned in Eiseley's passage. The extent to which the scientist in Eiseley's passage and the "some" in Belloc's passage agree or disagree is the manifestation of this contravercy.

In Loren Eiseley's passage their is a very concerned attitude expressed toward the relationship between science and human happiness. Eiseley views science as being destructive to human happiness through it's creation of atomic warfare. His protagonast replaces a turtle to it's natural habitat in a symbolic gesture of not further "tampering with the universe" as he had previously done as one of the architects of the atomic bomb. Eiseley obviously believes that science can go beyond a reasonable limit, that it can and has been a destructive force in human happiness. The very least which it has done is to cause one man (Eisley's protagonist) to question the morality of tampering with the universe in any way, even such an apparantly insignificant act as removing a tortoise from the wild. Eiseley's expressing through his protagonist his concern for the need to restrict the tampering with the universe and thus with human happiness. Science has the power to do this.

Hilaire Belloc expresses a somewhat different view of science than does Eiseley. Belloc observes the hardships in life and their relationship to ignorance. Science is knowledge; it discovers new things which conquor the ignorance and thus the hardships in human life. He places his foundation on the relationship between happiness and hardships, deducing that a decrease in hardships will produce an increase in human happiness. The validity of this foundation may be questionable, but Belloc uses it as a premise to indirectly saying that science has a direct bearing on human happiness. We need science, science will make us happier by removing our hardships. This is a very different view than Loren Eiseley's.

The extent to which the scientist in Eiseley's passage and the "some" in Belloc's agree or dissagree is the voicing of the contravercy between the need to promote or to limit science. The scientist and the "some" are not in agreement. Eiseley's scientist feels that science interferes with the universe, that it tampers with things which should be left alone. His tortoise is a very symbolic representative of this. Belloc's "some" feel that science promotes knowledge and ends hardship, thus helping people in their strive for happiness. This is a complete disagreement; a contravercy between the limits and morality of science.

Both Loren Eiseley and Hilaire Belloc side with their protagonists in their views concerning the necessary limitations of science. The argument between the scientist and the "some" is the same as between Eiseley and Belloc, both presenting different views on the limits and morality of science.

SCORE OF FOUR

In the above passages, Eiseley and Belloc seem to agree in most areas. It is my belief that Belloc's passage is written with sarcasm and that this is the factor which tends to unite the meanings of the two passages. I feel that both Eiseley and Belloc believe that science is ultimately destructive although it may seem to be a way to positive progress.

The final decision of the scientist, in Eiseley's passage to leave the turtle in its natural surroundings shows symbolically that the world should not be tampered with. Also, Eiseley demonstrates in his passage that the scientist has some regrets for creating something as destructive as the atomic bomb. I feel these are expressions of Eiseley's personal opinion on the subject of science and its negative effects upon nature.

I feel that Belloc, also, would agree that the scientist made the proper decision by letting the turtle remain in the woods. My reason for this opinion is that Belloc's use of the word "some" is outright sarcasm and is used to represent others besides himself. Yet another example of Belloc's sarcasm is the

part of the passage which says happiness comes from scientific discovery. Illustrating how far-fetched this idea is, the scientist in Eiseley's passage was not happy after his discovery of the atomic bomb but, instead, he was regretful.

Although scientific innovation seems to be a synonym for progress, I find myself forced to agree with Eiseley and Belloc. To me, science only tampers with nature in such a way as to worsen our natural conditions. Bellow and Eiseley, I'm sure, would be proud to here me state such an opinion.

SCORE OF FIVE

Throughout history, people have debated about the usefulness of science for the betterment of human life. Sometimes research threatened religion, and religious leaders raised voices in protest; sometimes scientific remedies were viewed as the path to human salvation.

These passages from Eiseley and Belloc illustrate the logical fear of the human hand interfering with nature. Eiseley relates a a story, almost a parable, about a designer of the worlds most devastating weapon, the atomic bomb. This physicist was overcome with joy at finding a tortoise to bring home to his children. But some small respect for nature or a pervasive guilt from creating the atom bomb causes him to put the creature back, as near to where he found it as possible because he fears that he has tampered with the universe enough for one man.

Of course, his major "tampering" has already been accomplished by the time he discovers his conscience. Eiseley is illustrating the human burden of responsibility that comes with scientific experimentation.

One doesn't know the effect that an invention can have on fellow human beings.

Invention of the machine did give liesure time to many, but also damned others to a life of drudgery and boredom, destroyed eardrums and bad backs. Creation of mechanized factories caused city crowding, which lead to crimes, pollution, and other urban problems.

Belloc, in capitolizing "Science", seems to banish the idea that "Science will bring happiness" to the realm of myth. The rhyme, simplistic and childlike, makes those who espouse this belief seem silly.

Belloc's "some", trusting and naive, rely on the wisdom of scientists, not realizing that they are human, prone to act on ego or for other human but less than humanitarian reasons. They equate intelligence and knowledge with wisdom, not realizing that scientists (or "Science"--"The Force") can have regrets like Eiseley's scientist coping with his guilt in a token way.

Both authors seem to agree with the scientist that humans mst must be cautious in their tampering and must think of all of the possible effects of what they are creating.

SCORE OF SIX

In the first passage the writer is illustrating by the anecdote that science is not necessarily a good or beneficial thing to human beings. The inventor of the atom bomb seems somewhat regretful that he used his scientific knowledge to invent something so destructive to people and the earth. This regret is illustrated by his reluctance to tamper with even the simplest laws of nature; that is the progress of the turtle in the woods.

The second passage takes a more ironic, mocking tone. It is saying the same thing as the first. By using a poem and the phrase, "some say". Who are "some"? The author makes us doubt seriously the quote from the "some" by deliberately leaving it vague as to what their credentials would be in making such a statement such as, "When Science has discovered something more, we shall be happier than before". This is very insubstantial and illogical. It does not follow that the cause of life being "hard and rough" is the failure of Science to "do more".

The "some" in the second passage don't agree with the scientist in the first. The "some" have no scientific knowledge; just a hazy idea that more scientific discoveries will make life easier. The scientist knows that this is not necessarily so because he has ample facts on which to base his knowledge. The development of the atomic bomb did not make life easier or less rough. In fact for a lot of people in Hiroshima there was no more life period because of this "scientific discovery". The "some" in the poem have no specific data to back up their claim--it is a generalized statement based on lack of specific knowledge of anything scientific.

The two authors seem to agree more with the scientist than the "some". The scientist in the first passage has found out what the "some" in the second passage still don't know. Science needs to be handled with care; all scientific progress, discoveries and work do not necessarily lead to health, wealth, and happiness. The "some" in the second passage are truly ignorant of this, and they are probably the very people who pay for scientists to work on anything bigger or more advanced in the hope that it will bring them "true" happiness.

The author of the second passage, by spotlighting this glaring ignorance, is mocking the "some", and is in agreement with the first author that science is not the panacea for all human ills.

TABLE CHART/READER ROSTER

June Essay Reading 1979

Question 1

Dr. Eileen Lothamer
Question Leader

Table 1

Lennis Dunlap
CSU, Chico
Richard Day
Humboldt State University
Stanley Cook
CSPU, Pomona
Jean Pickering
CSU, Fresno
Agnes Yamada
CSU, Dominguez Hills
William Walsh
CSU, Northridge

Table 2

Dan Buerger
San Jose State University
Peter Schroeder
CSC, San Bernardino
Lola Johnson
CSC, Stanislaus
Dan McLeod
San Diego State University.
Clifford Josephson
San Francisco State University
Clark Brown
CSU, Chico

Table 3

Larry Kramer
CSC, San Bernardino
John Cleman
CSU, Los Angeles
Marilyn Sutton
CSU, Dominguez Hills
John White
CSU, Fullerton
George Grant
San Jose State University
David Bell
CSU, Sacramento
Jake Fuchs
CSU, Hayward

Question 2

Dr. Rex Burbank
Question Leader

Table 4

Alice Brekke
CSU, Long Beach
Katherine Carter
CSU, Los Angeles
Michael Holland
CSU, Fullerton
Arthur Lane
CSU, Northridge
James Sullivan
CSU, Los Angeles
Solomon Iyasere
CSU, Bakersfield
Richard Hendrickson
Sonoma State University

Table 5

Tony Arthur
CSU, Northridge
David Kann
CPSU, San Luis Obispo
J. Barry Dalsant
Humboldt State University
Margaret Doane
CSC, San Bernardino
Lillian Wilds
CSPU, Pomona
Charles May
CSU, Long Beach

Table 6

David Rankin
CSU, Dominguez Hills
Hortense Thornton
CSU, Sacramento
Karen Hatch
CSU, Chico
Prescott Nichols
San Diego State University
Jack Conner
CSU, Hayward
John Harrington
CPSU, San Luis Obispo
James Frey
CSU, Fresno

STATISTICAL REPORT

August 1979

Robert Bradley[1]

I. The Test.

The seventh administration of the California State University English Equivalency Examination was conducted on Saturday, April 28, 1979.[2] The test has two ninety-minute sections; the first is an objective, multiple choice part, and the second is an essay part with two questions. The "Analysis and Interpretation of Literature" Subject Examination of the College-Level Examination Program (CLEP) was used as the objective part of the test. The two questions of the essay section were developed and scored by the English professors of the California State University and Colleges.

II. Scores.

1) Objective Score. The objective part of the test was scored by the Educational Testing Service, and scores were reported on the CLEP Subject Examination scale, which ranges from 20 to 80.

2) Essay Reading Scores. Each of the two essay questions was graded by two readers, and scores were given on a scale of 1 to 6. For a single reading of a question, a score of 4 or higher represented a passing performance. A zero indicated that the student did not attempt to answer the question.

3) Essay Question Score. The scores for the first and second readings of each question were combined to produce a question score, which ranged from 0 to 12.

4) Raw Essay Total Score. The two question scores were added to produce a raw essay total score. This score fell between the limits of 0 and 24.

5) Converted Essay Score. The raw essay total scores were transformed to the same scale as the objective scores to permit the addition of objective and essay scores in a meaningful way.

6) Composite Score. The objective score and the converted essay score for each person were added to produce a composite score. Inasmuch as both the objective and the converted essay scores are expressed on a 20 to 80 scale, the composite scores ranged from 40 to 160.

---

[1]I wish to thank Dr. John Bianchini, Head, Statistical Analysis Department, ETS Berkeley, for his contributions to prior editions of this report.

[2]See Cowell, W., The California State University English Equivalency Test, May, 1973. Statistical Report SR-74-10, Princeton, N. J.: Educational Testing Service, February 1974. See also Bradley, R., Statistical Report. In E. M. White, Comparison and Contrast 1978, Long Beach: Office of the Chancellor, the California State University and Colleges, 1978.

III. Candidate Groups and Norms Samples.

1) Total Group. The 3346 students who took both sections of the English Equivalency Test comprised the total group. The objective part of the test was taken by 3361 students, of which number 15 failed to take the essay section.

2) Equating Sample. The equating sample consisted of the first 523 papers scored at the essay reading session. All of the essay books from the various test centers were systematically scrambled prior to the reading to insure that they would be read in an essentially random order. Inasmuch as the means and standard deviations of both groups were not significantly different, the equating sample was assumed to be representative of the total group.

	Objective Scores	
	Mean	S. D.
Total Group	45.81	9.05
Equating Sample	45.14	8.86

The data from the equating sample were used to equate (transform) the essay scores to the objective score scale. A sample was used instead of the total group so that computer processing and statistical analysis could be done as the reading session progressed.

3) CLEP Norms Sample. This group consisted of 541 college students who participated in the 1964 norming administration of the CLEP Analysis and Interpretation of Literature examination.

4) Texas Norms Sample. This sample consisted of the 188 students included in the 1970 validity study of the CLEP examination at the University of Texas. Scores on the CLEP test were correlated with final grades in relevant courses.

5) CSUC Norms Samples. The essay questions of the English Equivalency Test were administered to two groups of California State University students. These essay papers, indistinguishable from regular papers, were interspersed with the candidate essays and scored during the reading session. The average essay scores for students who earned various letter grades in English composition courses were determined, and the information was used to help establish the minimum passing score for the essay part of the examination.

The larger of the two norms samples came from Humboldt State University, and consisted of 137 students in English 1, a 4-unit, one-quarter course of regular composition. The smaller sample was contributed by California Polytechnic State University, San Luis Obispo, and contained 60 students enrolled in English 114, a 4-unit, one-quarter course in composition, with special emphasis on expository writing. For both norms samples, the essay questions were administered to the students during the last several weeks of study designated for the course. The data from the two CSUC norms samples are summarized in Tables 1 and 2.

IV. The Equating Procedure.

Scores which are expressed on different scales cannot be combined without giving a greater weight to the scores on the scale which has the largest range of numerical values. The raw essay total score scale extends from 0 to 24, whereas the objective score scale extends from 20 to 80. If scores from these different scales were combined, the objective score would influence the composite score about three times as much as would the raw essay score. A solution to this problem is found by transforming the units of one of the scales so they are numerically equivalent to those of the other scale. This transformation can be performed so that the "equated scale units" represent equal increments in relative position on each of the score scales. Inasmuch as the objective scores were reported on the CLEP scale, the essay scores were transformed to the CLEP scale by means of the equipercentile method of equating. In so doing, a given converted score represents the same percentile rank on both the essay score distribution and the objective score distribution. In other words, equal converted scores on the two parts of the English Equivalency Test represent the same level of performance on each part.

V. Technical Notes on Equating.

1) In the strictest sense, the score transformation procedure was employed to establish scores that are comparable, rather than truly equivalent. The terms "equating" and "equivalent scores" are fully appropriate only when the two tests are parallel forms designed to measure the same psychological function. Because the objective and essay sections of the English Equivalency Test do not measure the same psychological function (see section XI), the score transformation technique produces scores which are comparable but not equivalent. However, the mechanics of transformation are exactly the same for producing either comparable or equivalent scores, and the only difference is one of interpretation. For convenience, the term "equating" is used rather than a technically correct phrase, such as "establishing comparable scores."[3]

2) Two basic methods of equating are available, and when the score distributions are very similar in shape, these methods yield almost identical results. Given similar score distributions, the linear method, which defines equivalent scores as those represented by equal standard-score deviates, is preferable. The linear method is completely analytical (determined by a mathematical relationship) and is free from judgments of curve smoothing. However, if the two score distributions differ in shape, the linear method fails to adjust for this difference.[4] Thus, when the distributions have different shapes, or when the shape of one or both distributions is not known or cannot be accurately predicted, the curvilinear or equipercentile method is to be preferred. If the two score distributions differ in shape, the equipercentile method will stretch or compress the score scale of the distribution being converted to fit the shape of the other distribution. This method defines equivalent scores as those scores which represent the same percentile rank on both distributions. Using the score conversion table (Table 3) for illustration, a raw essay score of 14 corresponds to a converted score of 45.

---

[3]Angoff, W. H. Scales, Norms, and Equivalent Scores. In R. L. Thorndike (Ed.), Educational Measurement, 2nd. ed., Washington, D. C.: American Council on Education, 1971, pp. 562-568; 590-597.

[4]ibid., p. 564

The percent of students scoring below 14 on the raw essay scale is the same as the percent scoring below 45 on the objective scale. Thus, in the sense that the two scores represent the same relative rank (percentile) within their distributions, they are equivalent scores.

3) Percentile rank values for each point on the objective score scale were plotted on normal probability graph paper. These points were connected by a smoothed curve, and for the points on the curve corresponding to the percentile rank of each essay score the objective score equivalent was determined.[5]

4) If the two score distributions differ in shape, the use of the equipercentile equating method, as noted above, will modify the essay raw score scale by altering the relative distances of some scores from the median. This modification of the raw score scale will result in a less-than-perfect correlation between raw scores and converted scores. Another factor can also contribute to an imperfect correlation. The equating process is based on a sample of essay scores, but the conversion process is applied to the total population of essay scores. If the shape of the total essay distribution differes slightly from the shape of the sample essay distribution, additional discrepancy may result between the raw score distribution and the converted score distribution. The correlation between raw essay and converted essay scores for the 1979 English Equivalency Test was .988. In prior years this value has always been above .99, and these very high correlations indicate that the linear equating method could have been used without inducing serious differences between the percentile ranks of equal scores on the converted essay and the objective score scales.

VI. Passing Scores.

1) Basic Passing Scores. An examinee was required to achieve a certain score on both the essay and objective parts of the examination in order to pass. These passing scores were added to arrive at a minimum composite passing score. Data from the CLEP norms sample were used to establish the passing score for the objective part. Because the average CLEP score for C students in both of these samples was 49, a score of 49 was set as the passing requirement on the objective part.

For the essay part of the examination, a score of 4 represented acceptable performance for a given essay reading. The criteria by which a reader was to distinguish between acceptable (4 or higher) and unacceptable (3 or lower) performance were carefully delineated in training sessions prior to the reading. These criteria represented a consensus among leading State University English professors. Each essay paper had two questions which were given two readings each, for a total of four reading scores. A score of 16 (on the 0 to 24 essay total score scale) thus represents an average score of 4 on each of the four readings. However, inspection of the frequency distribution of essay total scores reveals that only 22.6% scored 16 or higher, indicating that 16 would be a rather rigorous passing score. Also, setting the passing score at 16 would disallow any consideration for error in the scoring process. On the other hand, an essay score of 12 means that the average reading score was 3, or that on the average not one of the four readers found the performance on either of the two questions to be acceptable. A compromise was reached with a raw score of 14, which means that on the average two of the four readers judged the performance to be acceptable.

---

[5] Ibid., pp. 516, 571.

The basic passing score on the essay part of the exam was therefore tentatively set at 14.

The data from the CSUC norms samples (Tables 1 and 2) were also evaluated in an attempt to set an appropriate and defensible passing score for the essay part of the examination. The basic premise was that the essay passing score should be at least as high as the average essay score of students who earned grades of C in freshman English in the CSUC norms samples. For the Humboldt norms sample the average grade of C students was 12.89, and at San Luis Obispo the mean score for C students was 12.97. Inasmuch as students who pass the English Equivalency Test receive credit for two semesters of English, we might realistically expect their performance to be somewhat higher than that of C students in the norms samples, who had completed only one quarter of composition. It could be reasonably argued that the essay passing score should be at least as high as the average score of norms sample students who received a grade of C or higher (A + B + C). For the Humboldt sample this value was 13.70, and for San Luis Obispo it was 13.76. Thus a passing score of 14 represents a level of performance a bit higher than that achieved by "C or higher" students at either Humboldt or San Luis Obispo.

In addition, it should be noted that grades in freshman English classes may be affected by factors essentially unrelated to writing ability, such as attendance, promptness with assignments, attitudes, and motivation. Given these considerations, the grades earned by students in the norms samples were used only as supplemental evidence in setting the passing score, with primary weight being given to the essay scoring standards developed by the CSUC English professors. Accordingly, the decision to set the basic passing score at 14 represented a consensus among the project directors, chief readers, and consulting statisticians that this score was solidly supported primarily by the carefully developed standards of the CSUC English professors, and secondarily by the performance data in the norms samples.

2) Compensation Model. As in all prior years except 1977, a limited amount of compensation on essay and objective scores was permitted. That is, one of the scores could drop below the passing level if the other score was sufficiently high. Inasmuch as the basic objective passing score was set at 49, and that of the essay was set at 14 (which converted to 45), the composite passing score was set at 94 (49 + 45). The limits within which compensation would operate were then determined by defining scores for both parts which represented absolute minima, below which no compensation could be justified. The objective minimum was set at 45, because this score minimized the number of A, B, and C students who would fail and the number of D and F students who would pass in both the CLEP norms sample and the Texas norms sample.

The essay minimum score was set at 13, which means that at least one reader gave a satisfactory rating on one of the two essays. However, for a candidate to pass on the strength of only one acceptable reading it had to be accompanied by an objective score of 60 or higher, resulting in a composite score of at least 101. (In the CLEP norms sample, 60 was one point less than the average score of A students who had completed a year of study in literature, and represented the 85th percentile rank on the score distribution.) The limits of compensation, then, were as follows: essay scores of 13, which converted to 41, required an objective score of at least 60 with the resulting composite score of 101 or greater;

objective scores of 45 through 48 called for an essay score of at least 15 (which converted to 49) in order to yield the required minimum composite score of 94.

VII. Placement Scores.

This year, following a procedure initiated in 1977, a middle score category called "pass for placement" was designated. Students who did not pass for credit but who scored above a lower set of criteria were granted a recommendation for placement in a standard English composition course, and were also exempted from the English Placement Test (EPT). Students whose scores failed to meet these lower criteria received no recommendation and were required to take the EPT.

A simple conjunctive model was used for the "pass for placement" criteria. That is, no compensation was permitted, and a student had to achieve a certain minimum on both parts of the test. The essay score required for placement was set at 11 (36). A score of 11 can be represented by an average score of 2.75 on each of the four readings, or by three reading scores of 3 and one of 2. The project directors felt that a student who performed at this level exhibited the basic competence necessary for acceptable performance in a standard freshman composition class. The minimum objective score for placement was set at 35. This score represented a somewhat-better-than-chance performance on the multiple choice part of the test, and was indicative of at least minimal competence in reading interpretation.

In addition, all persons who received objective scores below 35 but whose essay scores were 15 (49) or higher were included in the "pass for placement" group. Such a relatively high performance on the essay would argue for the studen's readiness for a composition course, regardless of the low objective score. Fifteen students were added to the placement group on this basis.

VIII. Pass Rates.

1) Objective, Essay, and Total Pass Rates. A total of 3346 students took both parts of the English Equivalency Test. The 1081 students who passed the exam qualified to receive six semester units of credit in freshman English. The percentage of the total group which passed each part of the exam, using the basic pass criteria of 14 on the essay part and 49 on the objective part, is given below. The percentage passing the total exam without compensation (requiring 14 on the essay and 49 on the objective part) and the percentage passing with compensation (as defined in section VI, 2, above) are also listed.

	Objective Test		Essay		Total Exam			
					Without Compensation		With Compensation	
	No.	%	No.	%	No.	%	No.	%
PASS	1165	34.8	1587	47.4	866	25.9	1081	32.3
NO PASS	2181	65.2	1759	52.6	2480	74.1	2265	67.7

It is evident from these data that a score of 49 on the objective part represents a somewhat more difficult pass criterion than a score of 14 on the essay.

Accordingly, although only 25.7% of those students who passed the objective part failed the essay, 45.4% of those students who passed the essay failed the objective part. Expressed as percentages of the total group, 8.9% passed the objective test and failed the essay, and 21.5% passed the essay but failed the objective part. By contrast, in 1978 11.5% passed the objective test but failed the essay, and 19.9% passed the essay but failed the objective part. Thus in 1979 the disparity between the difficulty levels of the pass criteria was slightly greater than in 1978, with the objective pass criterion remaining somewhat more stringent than the essay criterion. On the other hand, prior to 1977 the pass criterion on the essay had been more rigorous than on the objective part.

2) The Effect of Compensation. The effect of the compensatory model was to raise the total pass rate from 25.9% to 32.3%. By scoring 15 or higher on the essay, 197 students with objective scores of 45 through 48 were permitted to pass. In addition, 18 students with essay scores of 13 but whose objective scores were 60 or higher were also allowed to pass.

IX. Technical Notes on Differences Among 1977, 1978, and 1979 Data.

1) Differences in Mean Objective Scores. The average objective score in 1979 (45.81) was 2.02 points lower than in 1978 (47.83) and 1.24 points lower than in 1977 (47.05). The lower mean score in 1979 is partly accounted for by the fact that alternate forms of the CLEP test are administered in successive years, and the more difficult form was used in 1979. Prior to 1978, the year-to-year variation in mean scores had consistently been about two points, with mean scores for every other year exhibiting only a very slight decline. These yearly variations in mean scores presumably reflected the difference in the difficulty level of the two tests. The slight drop in mean scores on the same form of the test used in alternate years apparently was due to a small yearly decline in the ability level (defined as the ability specifically measured by the objective test) of the test population. However, in 1978 the mean objective score was 1.17 points below the mean score in 1976, and in 1979 the mean score dropped 1.24 points below the 1977 level. These relatively large declines, demonstrated on both forms of the test, undoubtedly reflect a measurable drop in the ability of the test candidates in 1978 and 1979. Although this decline has not been large in an absolute sense, it has been a major factor in the rather sharp drop in the pass rate for 1979 (32.3%, down from 41.3% in 1978). The factor of lower candidate ability was also present in 1978, but in terms of effect on the pass rate it was masked by the fact that the easier form of the test was administered. In 1979, however, the use of the more difficult form of the test combined with the effect of lower candidate performance to produce a substantially lowered pass rate.

Although the two forms of the CLEP test have been equated, converted scores on one form of the test evidently do not represent exactly the same level of performance as identical scores on the other form of the test. This inference is based on the likelihood that the two forms may differentially emphasize the measurement of several complex intellectual functions. Thus the conversion resulting from an equating study would not be totally independent of the individuals from whom the data were drawn in developing the conversion, and such a conversion would not necessarily be applicable to all student populations. In summary, one form of the test could be easier than another for a particular student population.

2) Changes in Converted Score Equivalents. For the 1979 data, converted score equivalents for essay scores were slightly lower than those obtained in 1978. For

example, in the critical pass-fail range, essay scores of 13, 14, and 15 converted to 44, 48, and 52 in 1978, whereas in 1979 the same essay scores converted to 41, 45, and 49. Factors contributing to these differences in converted score equivalents include the changes in mean objective scores noted above, and year-to-year differences in the mean raw essay score. If the essay performance remains constant, a drop in objective scores results in lower converted score equivalents. Thus the effect of the somewhat lower objective scores in 1979 was to lower the converted score equivalents. This effect was partly counterbalanced, however, by the influence of a small decrease in the mean raw essay score in 1979. This slight drop in the essay scores means that a given essay score in 1979 represented a higher percentile than in 1978, and thus converted to a higher equivalent score on the CLEP (objective) scale. The net effect of the relatively large decrease in the mean objective score, partially offset by the slight drop in the mean essay score, was a three-point decline in the center of the scale.

X. Correlations Between Readings and Between Questions.

The correlation between the scores for the first and second readings of Question 1 was .70, and for Question 2 was .78. These figures, when expanded by the Spearman-Brown formula, represent the reliability of the grading procedure, and are discussed in the section on reliability. The scores for the first reading of Question 1 correlated .35 with the scores for the first reading of Question 2, and the scores for the second reading of both questions correlated .37.

When the two reading scores for each question are combined, the correlation between the questions scores is .42. This figure represents the degree of relationship between the two essay questions. Inasmuch as the proportion of shared variance is indicated by the square of the correlation, only 18% of the variance in each question was shared with the other. Thus suggests that the two questions drew upon essentially different types of writing ability. However, some of the underlying relationship between the two questions was obscured by the unreliability of the grading. An estimate of the true strength of this relationship can be obtained by removing the effect of the unreliability in the grading procedure. When the influence of grading unreliability is extracted from both questions by a procedure known as "correction for attenuation,"[6] the estimated relatonship between the questions is .50. The correction for grading unreliability increases the proportion of shared variance to only 25%, which tends to confirm the inference that the two questions represented different domains of content, and sampled different facets of writing ability which are presumably developed in English Composition classes.

XI. Correlations Between Parts.

Correlations among the objective, raw essay total, and converted essay scores were computed for the total group. The objective and converted essay scores correlated .59, which is about optimal for an examination of this kind. If the relationship were much lower it would be difficult to justify the compensation which is allowed by combining the scores to form a composite. A substantially higher correlation would mean that each part overlapped the other to the extent that the use of both parts would be redundant.

---

[6]Magnusson, D., Test Theory, Reading, Mass.: Addison-Wesley, 1967, pp. 147-149.

The correlation of .99 between the raw and converted essay scores indicates a strong linear relationship between the scores. As noted above (section V, 4), this high degree of linearity means that the two score distributions are very similar in shape, and the linear equating method could have been used without producing serious distortion.

XII. Reliability.

1) Reliability of the Objective Part. The reliability of the CLEP "Analysis and Interpretation of Literature" examination is .88 as reported in the CLEP Score Interpretation Guide.[7]

2) Reliability of the Essay Part. Because essay tests generally consist of only one or at most several items, it is not possible to compute reliability by split-half or internal consistency methods. The only alternative is "form-to-form" reliability, which entails giving two equivalent forms of a test to the same group of students. However, the preparation of an equivalent form for a one-or-two question test is not feasible, because one or two questions do not constitute an adequate sample from the domain of content. In lieu of a formal method of computing reliability, a method of estimation is available. Upper and lower limits of reliability can be established, with the certainty that the actual value lies somewhere between the two limits.

3) The Upper Limit. The correlation between grades on the first and second readings of each essay question represents the reliability of the grading procedure. Because unreliability due to grading is only of of several sources of error variance, the value for grading reliability will always be higher than the overall reliability value which included all sources of error variance. Thus, grading reliability can be used as the upper limit of overall or actual reliability. For Question 1, the correlation between grades on both readings is .70, and for Question 2 the correlation is .78. These values are the reliabilities of a single reading. That is, the reliability of each reading of Question 1 is .70. Adding the reading scores to obtain a question score produces an increase in reliability, which is estimated by the Spearman-Brown formula.[8] The estimated reliability for Question 1 is .82, and that for Question 2 is .87.

Because the essay total score is the sum of the question scores, the reading reliability for the total is higher than that for either of the questions. The reading reliability of the essay total score is most appropriately estimated by the formula for computing the reliability of a linear composite, inasmuch as the two questions (as established above) are not alternate forms representing the same domain of content.[9] The reading reliability of the essay total score is estimated to be .90, and this value is taken as the upper limit of reliability for the total essay test.

4) The Lower Limit. The obtained correlation between Question 1 and Question 2

---

[7] Score Interpretation Guide, College-Level Examination Program, Princeton: College Entrance Examination Board, 1967, p. 31.

[8] Magnusson, op. cit., p. 74.

[9] Nunnally, J., Psychometric Theory, New York: McGraw-Hill, 1967, pp. 226-230.

is used to estimate the lower limit of reliability. If each question were treated as a one-question test, the actual reliability per question could be obtained by giving an equivalent form of a question and correlating scores on both forms. Unreliability due both to grading and to fluctuations in performance would be reflected in such a "form-to-form" coefficient. It has already been demonstrated that Question 1 and Question 2 are not equivalent in terms of content. To the extent that the two questions are not equivalent, the correlation between them will underestimate the actual reliability of the question scores.

The correlation between the scores on Question 1 and Question 2 is .42. Again, because the essay total score includes both questions, the reliability of the total will be higher, and is estimated by the linear composite method to be .59. The actual reliability of the essay test therefore lies between the lower limit of .59 and the upper limit of .90.

5) Reliability of the Composite Scores. The estimated reliability of the composite scores, based on reliability coefficients of .88 for the objective test and between .59 and .90 for the essay part, is between .84 and .93.

6) Revised Policy on Discrepant Reading Scores and the Effect on Reliability. As noted in section II, each essay was read twice. And, as the scores for each reading ranged from 0 to 6, the discrepancy between the two reading scores could be as large as 5 points (scores of 0 and 6 could not be assigned to the same essay, as a 0 represents no attempt). In prior years, discrepancies of more than two points between reading scores were always resolved (reduced to two points or less) by a third reading. Two-point differences were allowed to stand, unless one reading was pass and the other was fail; that is, unless the scores were 5-3 or 4-2. In such cases, a third reading resolved the difference by taking the middle value for both scores or by assigning one of the values for both scores. In all cases involving a third reading, the third reader was not allowed to see the first two scores assigned.

Prior to 1978, the essay reliability figures published in the yearly "Statistical Report" did not reflect the resolution of the 5-3 and 4-2 score discrepancies; that is, the final scores assigned were slightly more reliable than indicated by the published coefficients. However, the reliability coefficients published in 1978 were based on the final scores assigned. To illustrate the effect of this more accurate procedure, prior to taking into account the resolution of the 5-3 and 4-2 score couplets for 1978, the correlation between the two readings of Question 1 was .58, and the corresponding value for Question 2 was .66. When the resolution of all 5-3 and 4-2 score pairs was included, these figures increased to .64 and .72. The actual reliability for each question involved the Spearman-Brown extension of the above values, as each question (essay) score was the sum of the two reading scores.

This year the project directors adopted a more restrictive policy regarding discrepant reading scores. In addition to the correction of all differences of three or more points, all differences of two points were resolved by a third reading (in contrast to correcting only two-point differences of 5-3 or 4-2). In effect, only one-point score discrepancies were allowed to stand. This new procedure resulted in slightly higher reliability coefficients than in previous years.

7) Reliability in Terms of Reading Score Discrepancies. Another means of expressing the reliability of essay scores is to indicate the frequency and magnitude of

the discrepancies between the two reading scores for each question. Inasmuch as each of the 3346 candidates wrote two essays and each essay was read twice, there were 3346 pairs of scores for each of the two questions. As noted above, the 1979 policy was to resolve all discrepancies of more than one point. Unfortunately, a small amount of error occurred in the discrepancy correction process, and 16 two-point discrepancies and one three-point discrepancy (out of 6692 score pairs) were not resolved.

For Question 1, 55.8% of the score pairs had no discrepancy. A discrepancy of one point occurred in 43.8% of the pairs, and a discrepancy of two points occurred (by error) in .4% of the cases. For Question 2, the figures were 59.5% for no difference, 40.4% for one-point differences, and .1% (by error) for two-point differences. In addition, a single three-point difference was mistakenly allowed to remain. In summary, for the total reading the two reading scores were either identical or differed by only one point in 99.75% of the cases, and two-point differences occurred inadvertently in .25% of the cases.

XIII. Validity.

1) Validity of the Objective Part. Course grades or other indicators of classroom achievement can be correlated with test scores to obtain measures of concurrent validity. The CLEP Score Interpretation Guide reports correlations between scores on the "Analysis and Interpretation of Literature" examination and course grades before and after the final examination; the values are .42 and .54, respectively. As noted in the CLEP Guide,[10] these correlations should be interpreted with caution because of lack of uniformity in the procedures by which the data were obtained.

2) Validity of the Essay Part. English composition grades earned by students in the CSUC norms samples were correlated with essay total scores. For the Humboldt sample, the essay/grade correlation was .24, and for the San Luis Obispo sample the correlation was .55. Inasmuch as class grades may be affected by factors unrelated to competence, such as absences or incomplete assignments, the instructors were also asked to rank their students on a six-point scale of general ability in the course. It was anticipated that the ranks would more accurately reflect the students' abilities than would the course grades, resulting in higher correlations with the essay scores. However, as in all prior years except 1978, the data were inconclusive. For Humboldt, the essay/rank correlation (.30) was higher than the value for essay/grade (.24), but for San Luis Obispo the essay/rank correlation (.50) was lower than the value for essay/grade (.55). Such relatively small differences between correlations have no substantive meaning, particularly in view of the rather small number of cases involved. It should also be noted that all the values cited in this section reflect unreliability in grading due to the various standards used by different instructors, and should be interpreted only as gross approximations, if not as underestimates, of concurrent validity.

TABLE 1

NORMS SAMPLE, HUMBOLDT

English 1, One Quarter

Grade	N	Mean	Standard Deviation	Range
A	16	14.56	3.57	10-23
B	83	13.88	2.46	9-21
C	36	12.89	2.49	9-20
D	1	7.0	0	0
F	1	12.0	0	0
A + B	99	13.99	2.67	9-23
A + B + C	135	13.70	2.67	9-23
C + D	37	12.73	2.64	7-20
D + F	2	9.50	2.50	7-12
TOTAL	137	13.63	2.72	7-23

50

TABLE 2

NORMS SAMPLE, SAN LUIS OBISPO

English 114, One Quarter

Grade	N	Mean	Standard Deviation	Range
A	2	20.0	2.99	17-23
B	13	14.62	2.44	11-19
C	30	12.97	2.63	9-19
D	12	11.25	1.74	8-15
F	3	10.67	2.06	8-13
A + B	15	15.33	3.11	11-23
A + B + C	45	13.76	3.01	9-23
C + D	42	12.48	2.53	8-19
D + F	15	11.13	1.82	8-15
TOTAL	60	13.10	2.99	8-23

TABLE 3

SCORE CONVERSION TABLE

June 1979

Essay Total Score	Converted Score
24	70
23	69
22	68
21	66
20	64
19	61
18	58
17	55
16	52
15	49
14	45
13	41
12	38
11	36
10	34
9	32
8	30
7	28
6	26
5	25
4	24
3	23
2	22
1	21
0	20

APPENDIX A

ENGLISH EQUIVALENCY EXAMINATION 1978-79

Start-Up Costs:

Personal Services	$18,298
Operating Expenses	13,346
TOTAL	31,644

Note: The above includes administrative time, secretarial and student assistant help, printing (brochure and essay test questions), mailing, and travel (including that of various faculty committees).

Reading Costs:

Readers and Professional Team	$17,640
Aides	1,800
Operating Costs	11,000
TOTAL	$30,440

Note: The above operating expenses include transportation and per diem.

Follow-up Costs:

Summer Administration	$ 3,000

Assistance underwritten by The California State University and Colleges campuses of San Jose, Pomona, and San Bernardino; and by the Office of the Chancellor.

Office Space, San Jose and San Bernardino
Telephone Usage, San Jose
Computer Time, Pomona
Key Punch Operator, Pomona
Statistician, Pomona
Computer Time for Follow-Up Study, Chancellor's Office
Development of Follow-Up Study, Chancellor's Office

APPENDIX B

THE ENGLISH EQUIVALENCY EXAMINATION: BUDGET SUMMARY

A Report to the CSUC English Council, Fall 1979

By Edward M. White

The budget for the EEE consists of two approximately equal parts: an appropriation from the New Program Development and Evaluation fund, and a student test fee. The fee (currently $20) is collected by campus test officers and sent to ETS for the College Board. The NPD&E grant (this year $67,649) is sent to the San Jose and San Bernardino campuses to cover the personnel and operating costs of administration, essay test development, and the June essay reading.

The $20 test fee covers three kinds of services we receive from ETS:

1. The development, construction, and printing of the College-Level Examination Program (CLEP) Subject Test, the Analysis and Interpretation of Literature. (This includes pre-testing, norming, and scaling.) That 90-minute multiple-choice test has been evaluated twice by English Council committees, which have recommended its use with some enthusiasm. The test is now being revised by a national committee, on which Professor Richard Lid, CSUN (and former associate director of EEE) serves. Test construction and revision costs for this kind of test approximate $100,000.

2. Test administration: Distributing test materials (including essay questions and booklets) to 20 test centers, dealing with test offices, paying proctors, printing a Test Manual. CSUC test offices receive $4-5 per test to cover proctoring and clerical costs.

3. Scoring and reporting services, including rosters, computer cards, and a standard statistical package. A separate transcript service is available for the CLEP portion of the test, which is widely accepted across the country. Statistical and essay reading consulting are available. ETS provides blank essay test books to CSUC specifications, and masking labels to conceal first scores from second readers.

The CSUC/NPD&E grant covers four kinds of expenses:

1. Personnel Services costs cover assigned time for White (.33) and Burbank (.375) and clerical support for them. 1979-80 total: $18,298

2. Operating Expenses cover supplies, travel, telephone, duplication, etc., as well as printing of brochures ($5,500), and essay test development ($2,200). 1979-80 total: $14,146

3. Essay Reading costs include honoraria, food, housing, and conference fee for all those involved in the June reading at Kellogg-West. 1979-80 total: $32,205

4. Summer program maintenance 1979 was $3000.

The following chart provides additional comparative information. Note that from 1973 to 1979-80, while the cost of living was doubling, the EEE budget has been trimmed by over 11%.

	1973	1977-78	1978-79	1979-80
EEE Budget from NPD&E (rounded)	$76,000	$75,770	$66,280	$67,650
Number of Examinations	4072	2826	3343	(3350) (est.)
State Per Diem (Based on cost of living)	$25	$40	$46	($50)[1]
Length of Essay Reading (in days)	4 1/2	4 1/2	3 1/2	3 1/2
Reader Honorarium	$300	$375	$325	($325)[2]
Reader Rate Per Day [3]	$67	$83	$93	($93)[2]

[1]Telephone "guesstimate" by state board of control.

[2]Project request for honorarium at $375 or approximately $107 a day was turned down, but negotiations continue.

[3]The contract with ETS for the English Placement Test was drawn up in accordance with the EEE reader rate for 1976, $325 for 4 1/2 days, or roughly $72 a day.

APPENDIX C

ENGLISH EQUIVALENCY EXAMINATION

DESCRIPTIVE STATISTICS - 1979 ADMINISTRATION

Division of Institutional Research
CSUC Office of the Chancellor
January 1980

The following tables present data describing the group of students who participated in the 1979 administration of the English Equivalency Examination. Tables are included which detail performance data and age characteristics of four groups of participants: (1) those who passed for unit credit; (2) those who were not awarded unit credit; (3) those who scored highly enough (including those who passed for credit) to satisfy the English Placement Test requirement; and (4) those who received neither unit nor placement credit.

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES

ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

TOTAL APPLICANTS BY AGE

Age	Non-Passers			Age	Placement Passers			Age	Credit Passers			Age	Total		
	N	Rel. %	Cum. %		N	Rel. %	Cum. %		N	Rel. %	Cum. %		N	Rel. %	Cum. %
16 & Under	1	0.2	0.2	16 & Under	6	0.4	0.4	16 & Under	2	0.2	0.2	16 & Under	9	0.3	0.3
17	27	4.9	5.1	17	76	4.4	4.8	17	76	7.0	7.2	17	179	5.4	5.6
18	378	68.4	73.4	18	1,271	74.4	79.2	18	727	67.4	74.6	18	2,376	71.1	76.8
19	72	13.0	86.4	19	150	8.8	88.0	19	98	9.1	83.7	19	320	9.6	86.3
20	7	1.3	87.7	20	17	1.0	89.0	20	23	2.1	85.8	20	47	1.4	87.8
21	4	0.7	88.4	21	13	0.8	89.8	21	11	1.0	86.8	21	28	0.8	88.6
22	5	0.9	89.3	22	12	0.7	90.5	22	8	0.7	87.6	22	25	0.7	89.3
23	1	0.2	89.5	23	6	0.4	90.8	23	11	1.0	88.6	23	18	0.5	89.9
24	1	0.2	89.7	24	4	0.2	91.0	24	7	0.6	89.2	24	12	0.4	90.2
25-29	7	1.3	91.0	25-29	15	0.9	91.9	25-29	35	3.2	92.5	25-29	57	1.7	91.9
30-34	4	0.7	91.7	30-34	5	0.3	92.2	30-34	11	1.0	93.5	30-34	20	0.6	92.5
35 & Over	46	8.3	100.0	35 & Over	133	7.8	100.0	35 & Over	70	6.5	100.0	35 & Over	249	7.5	100.0
All Ages	553	100.0	–	All Ages	1,708	100.0	–	All Ages	1,079	100.0	–	All Ages	3,340	100.0	–

TCSUC – Institutional Research

January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

MALE APPLICANTS BY AGE

Age	Non-Passers			Age	Placement Passers			Age	Credit Passers			Age	Total		
	N	Rel. %	Cum. %		N	Rel. %	Cum. %		N	Rel. %	Cum. %		N	Rel. %	Cum. %
16 & Under	–	–	–	16 & Under	1	0.2	0.2	16 & Under	–	–	–	16 & Under	1	0.1	0.1
17	7	2.7	2.7	17	24	3.9	4.0	17	23	7.4	7.4	17	54	4.5	4.6
18	189	73.8	76.6	18	475	76.2	80.3	18	210	68.0	75.4	18	874	73.6	78.2
19	43	16.8	93.4	19	73	11.7	92.0	19	37	12.0	87.4	19	153	12.9	91.1
20	2	0.8	94.1	20	8	1.3	93.3	20	4	1.3	88.7	20	14	1.2	92.3
21	2	0.8	94.9	21	8	1.3	94.5	21	3	1.0	89.6	21	13	1.1	93.4
22	3	1.2	96.1	22	5	0.8	95.3	22	1	0.3	90.0	22	9	0.8	94.1
23	1	0.4	96.5	23	3	0.5	95.8	23	4	1.3	91.3	23	8	0.7	94.8
24	1	0.4	96.9	24	3	0.5	96.3	24	3	1.0	92.2	24	7	0.6	95.4
25-29	4	1.6	98.4	25-29	4	0.6	97.0	25-29	14	4.5	96.8	25-29	22	1.9	97.2
30-34	1	0.4	98.8	30-34	1	0.2	97.1	30-34	4	1.3	98.1	30-34	6	0.1	97.7
35 & Over	3	1.2	100.0	35 & Over	18	2.9	100.0	35 & Over	6	1.9	100.0	35 & Over	27	2.3	100.0
All Ages	256	100.0	–	All Ages	623	100.0	–	All Ages	309	100.0	–	All Ages	1,188	100.0	–

Note: Sex data was unavailable for 188 applicants

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

FEMALE APPLICANTS BY AGE

Age	Non-Passers			Age	Placement Passers			Age	Credit Passers			Age	Total		
	N	Rel. %	Cum. %		N	Rel. %	Cum. %		N	Rel. %	Cum. %		N	Rel. %	Cum. %
16 & Under	1	0.4	0.4	16 & Under	5	0.5	0.5	16 & Under	2	0.3	0.3	16 & Under	8	0.4	0.4
17	20	7.8	8.2	17	52	5.3	5.8	17	53	7.3	7.6	17	125	6.4	6.8
18	188	73.4	81.6	18	793	80.6	86.4	18	512	70.7	78.3	18	1,493	76.0	82.8
19	29	11.3	93.0	19	77	7.8	94.2	19	60	8.3	86.6	19	166	8.5	91.2
20	5	2.0	94.9	20	9	0.9	95.1	20	19	2.6	89.2	20	33	1.7	92.9
21	2	0.8	95.7	21	5	0.5	95.6	21	8	1.1	90.3	21	15	0.8	93.7
22	2	0.8	96.5	22	7	0.7	96.3	22	7	1.0	91.3	22	16	0.8	94.5
23	–	–	–	23	3	0.3	96.6	23	7	1.0	92.3	23	10	0.5	95.0
24	–	–	–	24	1	0.1	96.7	24	4	0.6	92.8	24	5	0.3	95.3
25-29	3	1.2	97.7	25-29	11	1.1	97.9	25-29	21	2.9	95.7	25-29	35	1.8	97.0
30-34	3	1.2	98.8	30-34	4	0.4	98.3	30-34	7	1.0	96.7	30-34	14	0.7	97.8
35 & Over	3	1.2	100.0	35 & Over	17	1.7	100.0	35 & Over	24	3.3	100.0	35 & Over	44	2.2	100.0
All Ages	256	100.0	–	All Ages	984	100.0	–	All Ages	724	100.0	–	All Ages	1,964	100.0	–

Note: Sex data was unavailable for 188 applicants

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF CONVERTED ESSAY SCORES OF TOTAL PARTICIPANT POPULATION

Score	Frequency	Relative Percent	Cumulative Percent
20-24	13	0.4	0.4
25-29	54	1.6	2.0
30-34	325	9.7	11.7
35-39	843	25.2	36.9
40-44	524	15.7	52.6
45-49	831	24.8	77.4
50-54	286	8.5	86.0
55-59	312	9.3	95.3
60-64	115	3.4	98.7
65-69	39	1.2	99.9
70	4	0.1	100.0
Total	3,346	100.0	–

Mean = 43.8 Median = 41.3 Standard Deviation = 8.5

64

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF OBJECTIVE SCORES OF TOTAL PARTICIPANT POPULATION

Score	Frequency	Relative Percent	Cumulative Percent
22-24	5	0.1	0.1
25-29	61	1.8	2.0
30-34	245	7.3	9.3
35-39	539	16.0	25.3
40-44	731	21.7	47.0
45-49	683	20.3	67.4
50-54	500	14.9	82.2
55-59	309	9.2	91.4
60-64	188	5.6	97.0
65-69	79	2.4	99.4
70-74	17	0.5	99.9
75-78	4	0.1	100.0
Total	3,361	100.0	–

Mean = 45.8 Median = 45.0 Standard Deviation = 9.1

TCSUC – Institutional Research
January 1980

65

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF COMPOSITE SCORES OF TOTAL PARTICIPANT POPULATION

Score	Frequency	Relative Percent	Cumulative Percent
48-49	6	0.2	0.2
50-59	45	1.3	1.5
60-69	221	6.6	8.1
70-79	673	20.1	28.2
80-89	870	26.0	54.2
90-99	680	20.3	74.6
100-109	465	13.9	88.5
110-119	244	7.3	95.8
120-129	113	3.4	99.1
130-139	25	0.7	99.9
140-147	4	0.1	100.0
Total	3,346	100.0	–

Mean = 89.6 Median = 87.9 Standard Deviation = 15.7

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF CONVERTED ESSAY SCORES OF PARTICIPANTS WHO PASSED FOR CREDIT

Score	Frequency	Relative Percent	Cumulative Percent
41	18	1.7	1.7
45	162	15.0	16.7
49	257	23.8	40.4
52	223	20.6	61.1
55	159	14.7	75.8
58	113	10.5	86.2
61	69	6.4	92.6
64	38	3.5	96.1
66	26	2.4	98.5
68	10	0.9	99.4
69	2	0.2	99.6
70	4	0.4	100.0
Total	1,081	100.0	–

Mean = 52.7 Median = 52.0 Standard Deviation = 5.9

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF OBJECTIVE SCORES
OF PARTICIPANTS WHO PASSED FOR CREDIT

Score	Frequency	Relative Percent	Cumulative Percent
45-47	139	12.9	12.9
48-50	193	17.9	30.7
51-53	217	20.1	50.8
54-56	127	11.7	62.5
57-59	135	12.5	75.0
60-62	132	12.2	87.2
63-65	74	6.8	94.1
66-68	36	3.3	97.4
69-71	17	1.6	99.0
72-74	7	0.6	99.6
75-77	3	0.3	99.9
78	1	0.1	100.0
Total	1,081	100.0	–

Mean = 54.8 Median = 53.4 Standard Deviation = 6.4

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF COMPOSITE SCORES
OF PARTICIPANTS WHO PASSED FOR CREDIT

Score	Frequency	Relative Frequency	Cumulative Frequency
94	32	3.0	3.0
95-99	227	21.0	24.0
100-104	240	22.2	46.2
105-109	196	18.1	64.3
110-114	135	12.5	76.8
115-119	109	10.1	86.9
120-124	74	6.8	93.7
125-129	39	3.6	97.3
130-134	19	1.8	99.1
135-139	6	0.6	99.6
140-144	3	0.3	99.9
147	1	0.1	100.0
Total	1,081	100.0	–

Mean = 107.5 Median = 105.6 Standard Deviation = 9.7

TCSUC – Institutional Research
January 1980

69

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF CONVERTED ESSAY SCORES
OF PARTICIPANTS WHO DID NOT PASS FOR CREDIT

Score	Frequency	Relative Percent	Cumulative Percent
20-24	13	0.6	0.6
25-29	54	2.4	3.0
30-34	325	14.3	17.3
35-39	843	37.2	54.5
40-44	506	22.3	76.9
45-49	412	18.2	95.1
50-54	63	2.8	97.8
55-59	40	1.8	99.6
60-64	8	0.4	99.9
66	1	0.0	100.0
Total	2,265	100.0	–

Mean = 39.5 Median = 38.3 Standard Deviation = 5.9

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF OBJECTIVE SCORES
OF PARTICIPANTS WHO DID NOT PASS FOR CREDIT

Score	Frequency	Relative Percent	Cumulative Percent
22-24	5	0.2	0.2
25-29	59	2.6	2.8
30-34	242	10.7	13.5
35-39	537	23.7	37.2
40-44	727	32.1	69.3
45-49	441	19.5	88.8
50-54	162	7.2	95.9
55-59	74	3.3	99.2
60-64	15	0.7	99.9
65-69	2	0.1	99.9
70-73	1	0.0	100.0
Total	2,265	100.0	–

Mean = 81.1 Median = 81.8 Standard Deviation = 9.6

TCSUC – Institutional Research
January 1980

71

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF COMPOSITE SCORES
OF PARTICIPANTS WHO DID NOT PASS FOR CREDIT

Score	Frequency	Relative Percent	Cumulative Percent
48-49	6	0.3	0.3
50-54	16	0.7	1.0
55-59	29	1.3	2.3
60-64	58	2.6	4.8
65-69	163	7.2	12.0
70-74	273	12.1	24.1
75-79	400	17.7	41.7
80-84	423	18.7	60.4
85-89	447	19.7	80.1
90-94	319	14.1	94.2
95-99	102	4.5	98.7
100-104	23	1.0	99.7
105-109	6	0.3	100.0
Total	2,265	100.0	–

Mean = 81.1 Median = 81.8 Standard Deviation = 9.6

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF COVERTED ESSAY SCORES
OF PARTICIPANTS WHO PASSED FOR CREDIT OR PLACEMENT

Score	Frequency	Relative Percent	Cumulative Percent
36-39	734	26.3	26.3
40-44	499	17.9	44.2
45-49	810	29.0	73.2
50-54	282	10.1	83.3
55-59	309	11.1	94.3
60-64	115	4.1	98.5
65-69	39	1.4	99.9
70	4	0.1	100.0
Total	2,792	100.0	–

Mean = 45.7 Median = 44.8 Standard Deviation = 7.7

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF OBJECTIVE SCORES
OF PARTICIPANTS WHO PASSED FOR CREDIT OR PLACEMENT

Score	Frequency	Relative Percent	Cumulative Percent
35-39	425	15.2	15.2
40-44	660	23.6	38.9
45-49	649	23.2	62.1
50-54	476	17.0	79.2
55-59	298	10.7	89.8
60-64	185	6.6	96.5
65-69	79	2.8	99.3
70-74	16	0.6	99.9
75-78	4	0.1	100.0
Total	2,792	100.0	–

Mean = 47.8 Median = 46.8 Standard Deviation = 8.1

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF COMPOSITE SCORES
OF PARTICIPANTS WHO PASSED FOR CREDIT OR PLACEMENT

Score	Frequency	Relative Percent	Cumulative Percent
71-79	449	16.1	16.1
80-89	825	29.5	45.6
90-99	668	23.9	69.6
100-109	464	16.6	86.2
110-119	244	8.7	94.9
120-129	113	4.0	99.0
130-139	25	0.9	99.9
140-147	4	0.1	100.0
Total	2,792	100.0	–

Mean = 93.5 Median = 91.4 Standard Deviation = 13.8

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF CONVERTED ESSAY SCORES
OF PARTICIPANTS WHO DID NOT PASS FOR CREDIT OR PLACEMENT

Score	Frequency	Relative Percent	Cumulative Percent
20-24	13	2.3	2.3
25-29	54	9.7	12.1
30-34	325	58.7	70.8
35-39	109	19.7	90.4
40-44	25	4.5	94.9
45-49	21	3.8	98.7
50-54	4	0.7	99.5
55-58	3	0.5	100.0
Total	554	100.0	–

Mean = 34.0 Median = 33.9 Standard Deviation = 5.1

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF OBJECTIVE SCORES
OF PARTICIPANTS WHO DID NOT PASS FOR CREDIT

Score	Frequency	Relative Percent	Cumulative Percent
22-24	5	0.9	0.9
25-29	59	10.6	11.6
30-34	242	43.7	55.2
35-39	112	20.2	75.5
40-44	67	12.1	87.5
45-49	32	5.8	93.3
50-54	23	4.2	97.5
55-59	10	1.8	99.3
60-64	3	0.5	99.8
65-69	–	–	–
70-73	1	0.2	100.0
Total	554	100.0	–

Mean = 34.0 Median = 33.9 Standard Deviation = 5.1

TCSUC – Institutional Research
January 1980

77

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

DISTRIBUTION OF COMPOSITE SCORES
OF PARTICIPANTS WHO DID NOT PASS FOR CREDIT OR PLACEMENT

Score	Frequency	Relative Percent	Cumulative Percent
48-49	6	0.1	0.1
50-54	16	2.9	4.0
55-59	29	5.2	9.2
60-64	58	10.5	19.7
65-69	163	29.4	49.1
70-74	153	27.6	76.7
75-79	71	12.8	89.5
80-84	27	4.9	94.4
85-89	18	3.2	97.7
90-94	11	2.0	99.6
95-99	1	0.2	99.8
100-104	–	–	–
107	1	0.2	100.0
Total	554	100.0	–

Mean = 70.0 Median = 69.7 Standard Deviation = 8.2

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

INTERCORRELATIONS AMONG PART AND COMPOSITE SCORES
FOR THE TOTAL APPLICANT POPULATION

	Essay No. 1		Essay No. 2					
	Read 1	Read 2	Read 1	Read 2	Essay Total	Essay Conv.	Objective	Composite
Essay No. 1								
Read 1		0.7203	0.3882	0.4024	0.7518	0.7367	0.4442	0.6598
Read 2			0.4043	0.4138	0.7589	0.7443	0.4404	0.6619
Essay No. 2								
Read 1				0.7881	0.8094	0.8092	0.4714	0.7141
Read 2					0.8139	0.8124	0.4742	0.7179
Essay Total						0.9889	0.5878	0.8777
Essay Conv.							0.5879	0.8839
Objective								0.8979

TCSUC – Institutional Research
January 1980

THE CALIFORNIA STATE UNIVERSITY AND COLLEGES
ENGLISH EQUIVALENCY EXAMINATION – SPRING 1979

INTERCORRELATIONS AMONG GROUPED ESSAY AND COMPOSITE SCORES
FOR THE TOTAL APPLICANT POPULATION

	Question 1	Question 2	Essay Total	Conv. Essay	Objective	Composite
Question 1		0.4584	0.8198	0.8030	0.4769	0.7167
Question 2			0.8614	0.8601	0.5000	0.7596
Essay Total				0.9889	0.5878	0.8777
Converted Essay					0.5879	0.8839
Objective						0.8979

TCSUC – Institutional Research
January 1980